Economic Council
of Canada

Conseil économique
du Canada

Working Paper
Document de travail

92-05394/

Working Paper No. 34

The Influence of Socioeconomic Background on Education

Tom Siedule

1992

ISSN 1180-3487

The Influence of Socioeconomic Background on Education

The findings of this paper are the sole responsibility of the author and, as such, have not been endorsed by the Economic Council of Canada.

The Economic Council's Working Papers are protected by the Crown copyright and may be reasonably quoted under the terms of Canada's copyright laws. Requests for permission to reproduce or quote long excerpts should be addressed:

Director of Publications
Economic Council of Canada
P.O. Box 527
Ottawa, Ontario
K1P 5V6

Table of Contents

Acknowledgments vii

Foreword ix

Abstract xi

Introduction 1

Consensus and Disagreement 1

A Matter of Semantics 4

Historical Perspective 4

Methodology 8

Empirical Findings 12
- Eighteen-Year-Olds 13
  - Effects of Individual Factors 13
  - A Few Words of Caution 15
- Twenty-six-Year-Olds 15
- The Quantitative Effects of Socioeconomic Factors on Education 17

Concluding Remarks 20

Appendix 21

Notes 25

Figures

1 Years of education completed by 18-year-olds, Canada, 1986 5

2 Years of education completed and family income, 18-year-olds, Canada, 1986 6

3 Years of education completed by 26-year-olds, Canada, 1986 6
4 Wages and salaries of 26-year-olds, Canada, 1985 7
5 Weeks worked by 26-year-olds, Canada, 1985 7
6 Effects of socioeconomic forces on the education of 18-year-olds from rural areas, Canada, 1986 18
7 The effects of favourable socioeconomic factors on the education of 18-year-olds from rural areas, Canada, 1986 19

Acknowledgements

I would like to thank Keith Newton, Research Director of the Council's research project on the Canadian education system, for giving me the time and resources to develop them and to carry out the empirical work for this project. His comments at various stages have contributed greatly to what is good in this paper. All errors and omissions remain, of course, my sole responsibility.

Foreword

Since the publication of the 1966 Coleman report in the United States, policymakers, educators and economists have agreed that socioeconomic factors have a profound influence on education. This report summarizes an empirical investigation of this very important issue. Its specific contribution is the analysis of a large Canadian sample of individuals, using educational attainment (years of schooling) as the dependent variable. This contrasts with most studies of this sort, which have typically used smaller samples to evaluate young children's tests or examination scores — rather than attainment — as the outcome variable. The estimates here of the impact of a comprehensive set of individual socioeconomic factor on education will undoubtedly be of interest to policymakers. It is to be hoped that the results will also enhance public understanding of the issue.

This paper was prepared by Tom Siedule, an economist on the staff of the Council. It was undertaken as part of the Council's research program on the Canadian education system.

Judith Maxwell
Chairman

Abstract

The major concern of this study is whether socioeconomic status influences an individual's education. It uses the microdata file of the 1986 Population Census along with the analysis of variance technique, to determine this. The findings not only authenticate the correlation between educational attainment and socioeconomic status, but also quantify the effect of low socioeconomic status on an individual's education. The study estimates that the difference between very favourable and inordinately unfavourable socioeconomic influence can account for five years of additional education completed. The message is loud and clear. If the country is interested in raising the educational standard of all Canadians, then, in addition to the perennial search for a more efficient education system, policymakers have to find ways to alleviate some of the unfavourable social and economic influence on the less fortunate youths. The study estimates the overall influence as well as the factorial contributions of socioeconomic factors on education. It covers the population of all provinces and territories. The results should be relevant to both educators and social-policymakers.

Introduction

In 1966 Coleman *et. al.* concluded that the most important influence on educational attainment was socioeconomic background and not schools.[1] The report's famous claim that "schools make little difference to children's education" is still controversial, but researchers generally agree that, *a priori*, socioeconomic factors have a profound influence on education.

Policymakers are also aware of the impact of socioeconomic background on children's education. Project Head Start was a federally funded U.S. program designed to improve the educational achievement of children from disadvantaged socioeconomic backgrounds. Whether the project succeeded in achieving its goals has been a matter of controversy since its inception in 1965. However, a number of evaluations concur that the program had a significant impact on children from minority groups and on white children from single-mother families. The test scores of the program participants were consistently higher than the test scores of nonparticipants of similar socioeconomic status from the same schools.[2] Similarly, a benefit-cost analysis of the Perry Preschool Program finds that the return to investment in improving the education of the disadvantaged appears to be very high.[3]

Canadian evidence leads to a similar conclusion.[4] Unfortunately, Canadian studies are primarily based on the empirical findings of a single school board or a city. They are not representative of the whole country. If we are concerned with the educational attainment of all Canadian youths, then a study based on a nationwide coverage is a prerequisite for a better understanding of the topic. Moreover, most of the existing studies only provide policymakers with the correlation between socioeconomic status and children's educational attainment. They do not specifically estimate the influence of individual socioeconomic factors. The research reported here may help to fill this information gap.

Consensus and Disagreement

Project Head Start was the most famous policy measure implemented to offer disadvantaged preschool children special educational measures to compensate for background deficiencies. Even though the programs were also available to youths of higher socioeconomic status, the highest participation was from individuals of low socioeconomic status. Coleman *et. al.* concludes that Head Start programs tended to benefit only black children from the poor families of Southern states.[5] Subsequent research by others has found that the Head Start programs also benefitted

under-privileged youths of white and other ethnic origins.[6] Some researchers contend that the beneficial effects of the programs were transitory and the participants tended to lose the advantages gained after three years. In view of the fact that students stayed in Project Head Start for only a summer or a year, this debate over whether the effects were transitory or permanent seems immaterial. The question is whether or not the program benefitted the children. If the benefit lasted only one year after the children left the program, the program still demonstrated its effectiveness. Such an empirical finding simply suggests that low socioeconomic status has a powerful, detrimental impact on education. And for children of low socioeconomic status to excel (or simply to live up to their potential) in the education system, they need more than once-in-a-lifetime help at the beginning of their learning journey.

The Perry Preschool program, which was a very small project similar to Project Head Start, has also demonstrated beneficial effects on students and society. Weighing the program's benefits and costs to the society, one study estimates that the benefit-cost ratio of the program for society is six.[7] That is, for every dollar invested in the program, society would get back six in the long run. How is this possible? The answer is simple: higher tax revenue and a reduction in the cost of social services are more than enough to compensate for the investment. The benefit-cost analysis traces the participating and nonparticipating students' development from preschool days up to age 19. It includes data on educational attainment, employment, and criminal involvement. It concludes that a higher proportion of the participants were literate, employed, or enrolled in postsecondary education, and a lower proportion of the participants dropped out of school, were arrested for delinquent or criminal acts, or became welfare recipients.

In a sense, Project Head Start and the Perry Preschool program were the policymakers' answer to the socioeconomic problems of underprivileged children. In other words, policymakers presupposed the existence of the problems and used these programs to solve them. These programs, therefore, were not a direct test of the existence of socioeconomic influence on education.

Since the appearance of the Coleman report, economists have attempted to address the socioeconomic influence issue directly. All of them attempt either to prove the validity of Coleman's contention or to refute its conclusion. Most researchers contend that schools do matter in an extremely subtle way. For example, Murnane finds that the teacher's choice of techniques and general characteristics are important in predicting teaching effectiveness.[8] Summers and Wolfe contend that the effects of the school environment, facilities, and teachers can only be determined

from the details of microdata.[9] Specifically, they find that underachievers, students from low income families, and black students respond to some school inputs reasonably well. Murnane *et. al.*,[10] on the other hand, emphasize the importance of family background. They conclude that differences in the child care provided by mothers with different levels of education have a profound influence on children's achievements, but material wellbeing at home does not seem to have systematic impact on their education.

While there appears to be general agreement that socioeconomic factors significantly affect individuals' educational achievement, there are a number of methodological problems associated with these analyses. Most studies since the Coleman report focus on the effectiveness of schools, not the effectiveness of the individuals in the education system. This approach has certain pitfalls. For example, as a rule, the analytical unit has to be the school, and consequently, students' educational achievement is measured in terms of school averages. The aggregative nature of the data tends to disguise the school's true impact on individuals. Some studies[10] recognize the inadequacy of "school averages" and use student-specific microdata,[11] but they still fail to circumvent another serious problem: their geographical coverage is limited. It is methodologically difficult to derive a general conclusion from these studies, because what is true in, say, the Philadelphia school districts may not be so in other school districts. Some of these studies also use crude measures of socioeconomic influence (e.g., they use "black" or "white" to approximate racial differences and the student's family income to capture all socioeconomic influence) and leave out too many important dimensions of socioeconomic background.

The present report does not attempt to refute existing studies. Instead, it tries to give the reader an impression of the debate in the literature. There are some gaps. For example, a nationwide study on the influence of socioeconomic factors on education is overdue; furthermore, if the argument of the microdata users is valid, then only an analysis of a nationwide microdata base can give us an overall picture. The present study explores this relatively unknown area of education economics.

Ideally, a study of the output of the education system should cover the influence of institutional as well as personal attributes. Unfortunately, statistical agencies have found it impossible to integrate both dimensions in a nationwide micro database. The existing Canadian database includes information on individuals' educational attainment and their socioeconomic characteristics. The data are not, however, part of the school data. At present there is no acceptable method that enables researchers to link the information on individuals to the school data.

This study uses the available microdata from the 1986 Population Census to study Canadians' success and failure in the education system. Because of the limitations of the data, the study focuses only on the quantitative dimension of socioeconomic influence.

A Matter of Semantics

In education economics, many researchers use the same term to convey substantially different concepts. For example, the term "educational attainment" may mean different things to different authors. Some use the term to denote the grade-point average of a student, others use it to refer to the grade-point average of a class, a school, or a school board. Another example is the term "output of the educational system." This may denote the number of excellent, good, average, and inferior students produced by educational institutions. However, some analysts also use the term to denote variations in school attendance, student behaviour, or delinquency rates.

We do not dispute the importance of the many facets of education. In this study, however, the limitations of the data require us to use the term "educational attainment" to mean the number of years of formal education completed by individuals, and the term "socioeconomic factors" to denote social, economic, demographic, and geographic factors.

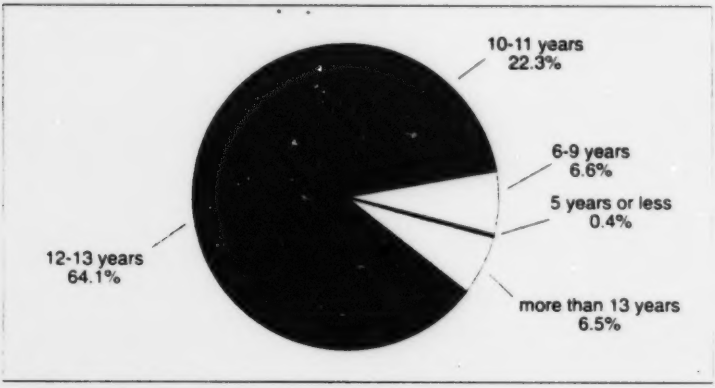
Historical Perspective

In 1986, more than 29 per cent of 18-year-olds in Canada completed at most 11 years of formal education, while 64 per cent completed 12 to 13 years, and more than 6 per cent of the youths of the same age group completed more than 13 years (Figure 1). "Years of education completed" refers to the level of education (grades and years completed) in elementary and secondary schools plus the total number of completed years of education at postsecondary institutions. If the individual took 14 years to complete grade 12, he/she would have only 12 years of education completed. On the other hand, a youth who finished grade 13 in eleven years would have the credit of 13 years of formal education.

The statistics show that there are some "over achievers" as well as many "under achievers." In Canada, children usually enter grade one at the age of six. Therefore, by the age of 18, they should have completed more than 11 years of formal education. The question is why did one in three 18-year-olds not reach their potential in the education system in 1986?

Figure 1

Years of education completed by 18-year-olds, Canada, 1986



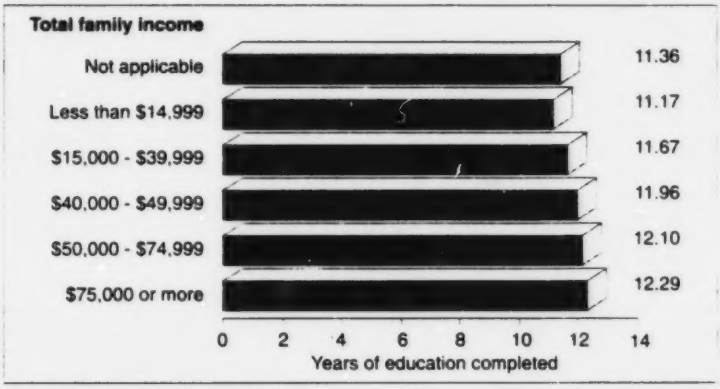
SOURCE Based on data from the 1986 Population Census.

A quick examination of the available data would reveal that economic conditions (approximated by family income) had an effect on the average educational attainment of 18-year-olds (Figure 2). Youths from the higher-income families had, on average, more years of education than their counterparts from lower-income families. From a methodological point of view, however, there are reasons to question the implications drawn from the raw data. First, is the difference large enough to be statistically significant? Second, family income is only one of the many factors that may influence the level of educational attainment. For example, ethnic origin, whether the setting is urban or rural, the residential environment, the province of residence, the size of the family, and whether or not the individual is from a husband/wife family may also have had an important impact on individual educational attainment. Researchers must systematically test these hypotheses with the appropriate data and statistical techniques.

In terms of educational accomplishment, the 26-year-olds were similar to the 18-year-olds — 27 per cent had completed 11 or fewer years of formal education and 9 per cent failed to reach grade 10 (Figure 3). For 26-year-olds, the measurement of success must go beyond the education system and include experience in the labour market. The majority of individuals in this age group should have completed their formal education by 1985. Wages and salaries and weeks worked in a year should, therefore, be good measures of their performance in the labour market. Apparently, a sizable segment of them (25-year-olds in 1985) did not earn a good living in 1985. A total of 46 per cent made less than $10,000 in

Figure 2

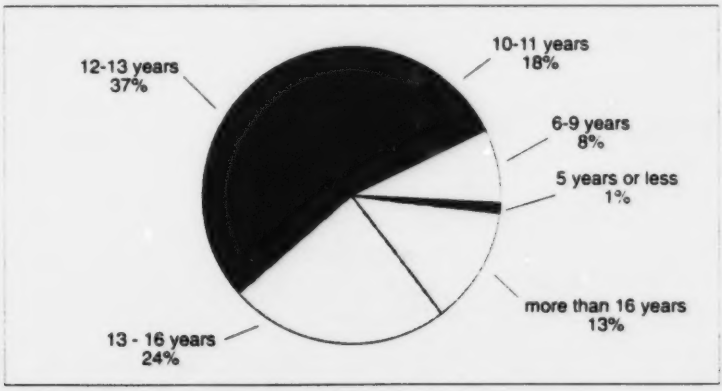
Years of education completed and family income, 18-year-olds, Canada, 1986



SOURCE Based on data from the 1986 Population Census.

Figure 3

Years of education completed by 26-year-olds, Canada, 1986

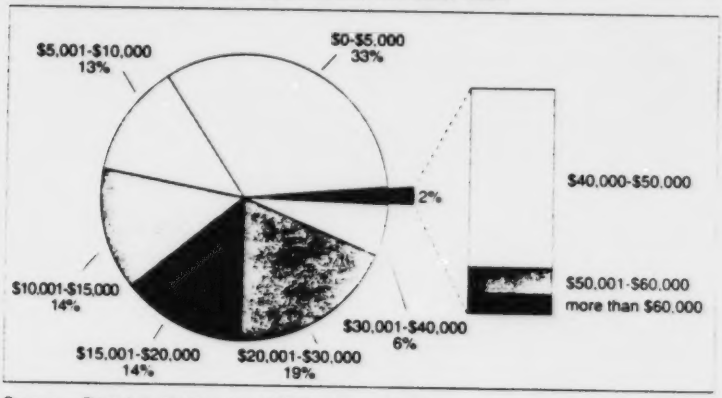


SOURCE Based on data from the 1986 Population Census.

1985; an overwhelming 33 per cent had annual incomes of $5,000 or less (Figure 4). Statistics on weeks worked reveal a pattern that is consistent with the wage-and-salary information. Thirty-five per cent of the people in this age group worked less than 30 weeks in 1985, including a majority who had gainful employment for less than 10 weeks in 1985 (Figure 5).

Figure 4

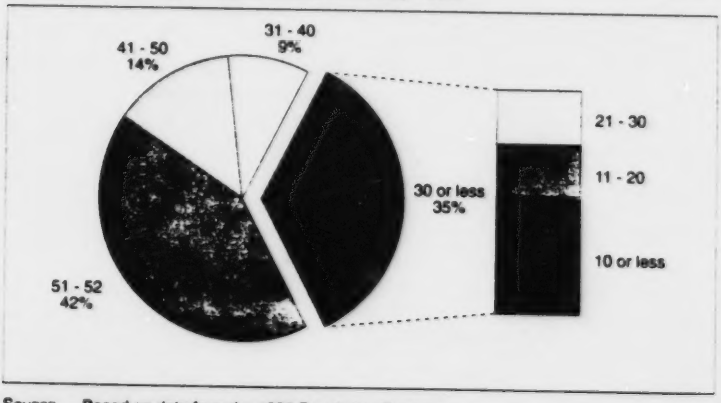
Wages and salaries of 26-year-olds, Canada, 1985



SOURCE Based on data from the 1986 Population Census.

Figure 5

Weeks worked by 26-year-olds, Canada, 1985



SOURCE Based on data from the 1986 Population Census.

Are low educational attainment, low employment income, and involuntary unemployment interrelated problems? Can they be traced to the same socioeconomic factors? These are some of the questions that the present paper hopes to answer.

Methodology

For any empirical study, the researcher has to work with what is available in the database. Since most of the data for this study is drawn from the 1986 Population Census, the empirical work has to be defined in terms of what is in the census. For this reason, the number of years of formal education completed is taken as the dependent variable for 15–22 year-olds. For individuals aged 23–26, the focus is broadened to include three dependent variables — years of formal education completed, wages and salaries earned in a year, and the number of weeks worked in a year — in an attempt to determine what "success" means to more mature individuals at this stage of their career. The census contains sufficient information for researchers to calculate all three dependent variables. To calculate number of years of formal education, it contains information on the highest degree, certificate, or diploma obtained; the highest grade of elementary or secondary school completed; the total number of years completed at a degree-granting educational institution; and the total number of years completed at a non-degree-granting educational institution. As well, it includes weeks worked and wages and salaries earned in 1985.

The census also contains information on a number of potential explanatory (independent) variables, such as age, gender, province of residence, whether the residence is urban or rural, census family status and living arrangement, ethnic origin, whether the person worked full time or part time in 1985, hours worked in the reference week, labour force activity, major field of study, number of persons in a census family, total census family income, and the value of the dwelling. Obviously many of these are classification variables. For example, "province of residence" is by definition not a metric measure, and must be coded by a classification name or a discrete number. Even total family income in the census microdata is coded by income intervals, not by exactly how much the family made in 1985. Fortunately, classification independent variables are perfectly suitable for the statistical analysis of variance (ANOVA).

We hypothesize that an individual's educational attainment depends on six socioeconomic factors: family income, commercial value of dwelling, ethnic origin, province of residence, family size, and family status/living arrangement. Family income as a socioeconomic factor is self explanatory, but the influence of the remaining independent variables requires elaboration. The commercial value of the dwelling where the child lives is a proxy for the environmental effects of the location of the residence and school. With the exception of the children of wealthy families who are sent to private schools, Canadian youths usually attend schools in the neighbourhood of their residence. Naturally, a well-to-do

neighbourhood tends to be associated with a more favourable environment for academic endeavour, whereas a school in a "ghetto" is likely to encounter more problems with drug abuse and delinquency. The influence of ethnic origin on children's educational attainment is empirically verified by Cheng *et. al.*[12] The effects of the province of residence variable is less obvious. They may come from two sources. First, cultural and circumstantial differences among the provinces influence an individual's perception of the optimal amount of education necessary for his/her work (e.g., a farmer from the prairies would probably not consider a Ph.D. essential for raising livestock). Second, provinces do not devote identical amounts of resources to education (the federal government's equalization payments on education help, but they cannot really put all the provinces on an equal footing). These differences among provinces should, therefore be manifested in the educational attainment of their populations. "Family size" affects children's educational attainment in a particular way.[13] *Ceteris paribus*, we expect a threshold value for the family size to be most favourable for educational accomplishment. A youth from a one-child family does not have any siblings with whom to share academic experiences, which may work against the child's educational aspirations. On the other hand, if a family has too many children, those children have to share not only the limited financial resources of the family but also their parents' limited time for guidance. Finally, the family-status variable (whether the child lives in a husband/wife household, a single-parent family, or with relatives or unrelated individuals) is self explanatory.

Although the quantitative effects of these variables have to be empirically verified, few researchers would reject them as potential explanatory variables. Conceptually, however, there are other socioeconomic variables that fit the hypothesis well. The list given here represents the variables that enter the finalized version of the statistical exercises. The study excludes other variables for various reasons: (1) they are not included in the 2-per-cent sample of the 1986 census tape; (2) they were initially included but subsequently dropped because they showed no statistically significant influence on a person's educational attainment; or (3) they are not essential to the socioeconomic theme.

Conceptually, there is nothing to prevent the researcher from specifying the analysis outlined above in a multiple regression. Practically, however, there are a number of problems with this approach. For example, the researcher may set the dependent variable equal to the "years of formal education completed" and the independent variables equal to the socioeconomic attributes associated with the individual, which may be expressed as large sets of dummy variables in terms of "0" and "1." The dependent variable presents no problem, since the child's educational

attainment is measured in terms of number of years of education completed. The independent variables are, however, more complex. The census microdata file presents family income in terms of 24 income groups. The province-of-residence variable is similar. The child must have come from one of the 10 provinces or the Yukon/Northwest Territories. This means that for the province of residence variable alone, the researcher needs 10 dummy variables to capture the provincial effects on education.[14] In addition to five sets of dummy variables to represent five socioeconomic attributes, the researcher has to consider the interaction effects. For example, the average family income is much higher in Ontario than it is in Saskatchewan. Therefore, we may expect a family income of $50,000 per year to have different influence on the educational attainment of the individuals in these two provinces, which is technically the interaction effect of the family-income and province-of-residence variables. Similarly, the researcher may consider the interaction effects of other socioeconomic attributes. The finalized estimating equation could easily have more than 150 dummy variables and interaction terms. In applied econometrics, an equation with so many regressors (independent variables) is generally considered unwieldy. However, there are other reasons for not applying the multiple regression technique to the present problem. For example, if the influence of the province of residence on the child's education is not different from the national average, then the estimated coefficient for this dummy variable would turn out to be statistically insignificant. On the other hand, the influence of selected provinces can be substantially different from the national average and the estimated coefficients for these dummy variables can be highly significant. Thus, the estimated coefficients provide us with certain indications of provincial effects individually. However, we still have no clear answer to the simple question. "Does the province of residence make any difference to a child's education?" The researcher can only answer this question by testing the statistical significance of the *whole set* of coefficients for the provincial dummies once again (i.e., performing the test after estimating the equation). Moreover, the researcher occasionally wants to know whether or not the effect of one region (e.g., the Atlantic region) is significantly different from that of another region (e.g., Ontario). Without additional statistical calculations, the estimated equation does not give the researcher this information. Thus, in practice, forcing the socioeconomic hypothesis to fit the "straight jacket" of multiple regression is not the best statistical method.

Analysis of variance (ANOVA) is a statistical method that can be used to study the effects of multiple independent variables (all in classification form) on one or more dependent variables (all in metric measure). When the researcher relates one dependent variable to a set of independent variables, it becomes an application of univariate analysis of variance. When

the researcher attempts to measure the effects of multiple independent variables on two or more dependent variables simultaneously, it becomes a problem of multivariate analysis of variance (MANOVA). In this study, we have applied the data of individuals aged 15–22 to univariate analysis of variance. The data for individuals aged 23–26 were subjected to both univariate and multivariate analyses, because for the study analyzes not only their educational attainment but also their employment income and weeks worked. The empirical results of ANOVA and MANOVA are consistent with each other, but we have found that the MANOVA findings shed no further light on the issue. For this reason this report does not discuss the MANOVA results any further. The remainder of this report uses the term analysis of variance, or ANOVA, strictly in its univariate analysis context.[15]

The socioeconomic hypothesis outlined above is a typical of the kind of problem to which researchers apply ANOVA. Technically, the *null hypothesis* is that socioeconomic factors have no effect on individual educational attainment. Correspondingly, the alternative to the *null hypothesis* is that socioeconomic factors do matter. We propose to include six socioeconomic factors in the test. Each factor (independent variable) has several classes (e.g., the province-of-residence variable has 11 classes) and a factor may interact with other factors. The statistical technique, specified in terms of a high-order structure-of-variance analysis, is capable of taking into account the effects of all socioeconomic factors simultaneously.

For the statistical analysis, in addition to the necessity of having all dependent variables in metric measure and independent variables in classification form, the researcher must know the characteristics of the data; i.e., are they balanced or unbalanced? The ANOVA statistical technique for *balanced data* is quite different from that for unbalanced data.[16] Applying the inappropriate technique to the unbalanced data can lead to misleading conclusions. Unfortunately, in economic data, unbalanced design is the norm rather than an exception. For example, if the researcher hypothesizes that youths from the middle and high income families tend to have higher educational attainment than those from low income families, then the researcher has a simple one-way ANOVA problem. However, this says nothing about balanced and unbalanced designs. The researcher has to learn this from counting the number of individuals in each cell of the ANOVA design. If the number of youths from low income families is not equal to the number of youths from middle and high income families, then the problem automatically becomes a one-way ANOVA of unbalanced design.[17]

The Generalized Linear Model of ANOVA allows us to estimate the expected educational attainment of an individual, subject to the influence

of a specific combination of socioeconomic factors (e.g., a combination of many unfavourable factors or a combination of the most favourable factors). It can be performed as an option in the statistical run of ANOVA. It not only produces data on the estimated years of education completed for the individuals in question, but also tells us whether or not the estimate is statistically significant. In short, the analysis of variance is a more convenient and powerful technique than traditional multiple regression for the empirical work of this study, permitting as it does the analysis of metric (dependent) variables and multiple classification (independent) variables and providing a useful set of statistics and estimated equations of coefficients (the estimated equations are available in the Generalized Linear Model of ANOVA).

For descriptive convenience, this report primarily describes and discusses the historical data and statistical results for individuals aged 18 and 26. Data from all the provinces and territories are included. The coverage and the technique of applying the analysis of variance to microdata make this study, perhaps, the first of its kind in Canadian education research.

Empirical Findings

A priori, young people from urban centres have different educational aspirations than do those from rural areas. For this reason, the empirical analysis classifies all individuals aged 15–26 according to whether their residence is urban or rural. In other words, we analyze the urban and rural subpopulations separately. The two subpopulations are: (a) urban — individuals who lived in census metropolitan areas with urbanized cores of 100,000 or more population in 1986, and (b) rural — all other individuals who lived in small cities, towns, and rural areas. This data stratification automatically separates the effect of urban environment on education from that of rural setting. The task of the ANOVA exercises is, then, to determine the influence of other factors on education.

This study recognizes the relationship between formal education and age. The statistical work analyzes the data pertaining to individuals aged 18 and 26 separately. These ages represent two critical stages of young people's educational and career development. At age 18 an individual typically has to decide whether or not to enter the labour force for the first time or to pursue further education and training. At age 26 a person has usually completed his/her education and may have even been in the work force for a few years already. Statistical analyses of these two age groups are, therefore, essential to the socioeconomic theme. The results of these statistical analyses for those aged 18 and 26, therefore, are reported in Tables A-1 to A-8.[18]

Eighteen-Year-Olds

As shown in Tables A-1 and A-2, the factors included to explain the educational attainment of 18-year-olds from urban and rural areas are highly significant by usual statistical criteria. The statistics are for Type III analysis of the Generalized Linear Model of ANOVA. In this type of analysis, the main effect of a factor (independent variable) is adjusted for the effects of other factors and interaction effects. The statistical analysis also gives us the results of other types of ANOVA analyses, which do not fully take the adjusted effects of Type III into consideration.[18] These results are not reported here. Suffice it to point out that the statistical results for the other types of analysis are at least as strong as the ones reported in Tables A-1 and A-2. These tables show that, in addition to the obvious differences in statistical significance, ethnic origin appears to have significant influence on the educational attainment of 18-year-olds from rural areas but not on that of youths from large urban centres. Perhaps the melting pot of large urban centres assimilates youths of different ethnic origins and dilutes the effect of cultural background. The corollary of this argument is that youths of different ethnic origins from rural areas do not have sufficient opportunity to meet people of other cultural backgrounds. Therefore, they tend to preserve the cultural flavour of their ethnic origins. The interaction terms for income class/province of residence and income class/family size also turn out to be statistically significant for youths from large urban areas but insignificant for 18-year-olds from rural areas. Since income class, province of residence, and family size by themselves have a significant influence on the educational attainment of 18-year-olds in both urban and rural settings, the interaction terms capture the effects over and above those of the three variables separately. The interaction effects conceptually show the desirable or detrimental influence of the combination of socioeconomic attributes.

Effects of Individual Factors

Family income — For 18-year-olds from both large census metropolitan centres and rural areas (with the exception of the "not applicable" class which covered youths in collectives, in households outside Canada, temporary residents and nonfamily persons), youths from well-to-do families consistently completed more formal education than did those from poor families. The educational attainment of youths from families with incomes of $40,000–$49,999 or greater is statistically higher than that of individuals from families with annual incomes of less than $14,999. The influence of family income on education is unambiguous.

Value of dwelling — This variable is highly significant, but its effect on education requires careful interpretation. Youths living in rented households tend to have less education than those living in family-owned houses. The educational attainment of youths from expensive dwellings is, however, only slightly higher than that of those from less expensive residences. The ANOVA Contrast Test confirms that the environmental difference between family-owned houses and rented houses is responsible for the statistical significance of this variable. Therefore, for both the urban and rural areas, this variable captures the different effects of rented and owned houses rather than the influence of property values on educational aspiration.

Province of residence — The province-of-residence variable significantly affects the educational attainment of youths from large urban centres and rural areas, but the statistical significance is weaker for urban than for rural areas. Youths from large urban centres in Quebec and Ontario slightly outperform their counterparts in Manitoba, Saskatchewan, Alberta, and British Columbia. (It should be noted that the 2-percent sample 1986 census tape does not include a single census metropolitan area in the Atlantic provinces, the Yukon, or the Northwest Territories. Therefore, the urban-population sample does not cover individuals from these provinces and territories.) The educational attainment of youths from rural areas exhibits more variation across provinces. Generally speaking, 18-year-olds from the Atlantic provinces completed less education than those from other provinces. Youths from the Yukon and the Northwest Territories tend to have the least formal education.

Ethnic origin — The 1986 census classifies individuals into 28 ethnic groups. However, if the data are cross-classified according to the urban/rural settings and the 28 ethnic origins, many cells contain few if any observations. For this reason, the 28 ethnic groups have been aggregated into five large groups. However, while this procedure circumvents the data problem, it creates a special problem for the group labelled "minority," which now consists of ethnic individuals of heterogeneous origin, including those of West Asian and Arab origin, South Asians, Chinese, Filipinos, East and South-East Asians, aboriginal people, blacks, African blacks, and those from the Caribbean. Because of this aggregation, the statistical analysis fails to detect the diversity among people of minority-group origin. Given this data limitation, the ethnic-origin variable is only significant for the results from the rural areas.

Family size — In both urban and rural areas, youths from four-person families tend to have more formal education than those from larger and smaller households. Youths from families (regardless of family size) tend

to do better than those of the same age from nonfamily households. The overall impression is that a family of four people is most favourable for children's pursuit of formal education, while very large and very small families and the nonfamily environment seem less conducive to academic achievement.

Family status and living arrangement — Children from husband/wife households and single-male-parent families have a better chance of completing more formal education than those who live with single mothers, relatives, nonrelatives, common-law spouses, or alone. This underlines the importance of the home environment on education. This is true of urban and rural areas.

A Few Words of Caution

The associations drawn between educational attainment and socioeconomic factors imply extremely complex relationships. Socioeconomic factors can influence but cannot cause an individual's success or failure in the education system. In the real world, family income approximates several factors. In some cases it is a proxy for the parents' educational background, the parents' ability to organize and use their time properly, and their ability to guide their children through the "jungle" of the education system. If this is true, then it is the implicit quality of the family income rather than the dollar value of the income that influences a child's education. Similarly, it is the environment created by the value of dwelling, family size and family status/living arrangement that contributes to individuals' success or failure in the system. The upshot is that policymakers have few feasible options in attempting to mitigate the effects of low socioeconomic status. High educational attainment requires proper educational facilities and favourable environments for all Canadian youths. So far, neither the government nor educators have been able to provide less fortunate youths with these amenities adequately.

Twenty-six-Year-Olds

As mentioned earlier, for 26-year-olds the study attempts to determine success in life, which includes both success in the education system and in the labour market. Accordingly, the research involves pinpointing the socioeconomic factors that can significantly affect individuals' years of education completed, employment income, and weeks worked in a year. Since educational achievement can affect an individual's income and employment, and weeks worked has a direct impact on employment income, new classification variables have to be created when used as

explanatory variables. In particular, the research design regroups the years of education completed into three education classes. Individuals with less than 10 years of education are grouped into education class 1, those with more than 11 but less than 12 years of education are grouped into education class 2, and those with 13 or more years of formal education are grouped into education class 3. Weeks worked are similarly coded into a classification variable where needed. Tables 3 to 8 summarize the results for 26-year-olds from urban and rural areas and for each of the dependent variables separately.

The results for the educational attainment variable are consistent with those reported for 18-year-olds. Family income, value of dwelling, ethnic origin, and family size show significant influence on the educational attainment of young people aged 26 from both urban and rural areas. Provincial differences contribute to the educational achievement of 26-year-olds from rural areas in more than one way. First, the province-of-residence variable by itself affects a 26-year-old's chance of success in the education system. Second, province of residence also interacts with family income to exert further influence on an individual's education. For 26-year-olds from census metropolitan areas with populations of 100,000 or more, the interaction between province of residence and value of dwelling has an indirect effect on the individuals' education, but the impact of province of residence by itself is insignificant.

Although the results for 26-year-olds appear to be consistent with the findings for 18-year-olds, the implications are not necessarily identical. The statistical associations between the educational attainment and family income variables are for the two ages similar. As expected, individuals from well-to-do families tend to complete more education. However, family income does not have the same connotation for 18-year-olds as it does for 26-year-olds in 1986. The majority of 18-year-olds are still full-time students and do not contribute to family income. On the other hand, the majority of 26-year-olds are already members of the labour force and are probably major contributors to family income. Therefore, the correlation between their educational attainment and family income takes on different meaning. The relationship in this case is a statistical association (i.e., educational attainment and family income tend to move together) with no clear indications of cause and effect. Because most individuals aged 26 have already completed their formal education, the empirical relationships between the socioeconomic factors and educational attainment must be interpreted with discretion.

The results show that province of residence, family size, education and weeks worked have a statistically significant influence on employment income. The impact of province of residence, years of education

completed, and weeks worked, on employment income is clear. The correlation between family size and employment income, however, is probably a statistical association rather than a cause-and-effect relationship.

The statistical results illustrate that many factors — including family income, province of residence, family size, education class, and the interaction effect between family income and education class — systematically affect individuals' work experiences (i.e., weeks worked). This is true of individuals from large census metropolitan areas as well as those from rural areas. The relationship between work experience and education is such that, *ceteris paribus*, more educated individuals always have more success in the work place than less educated individuals.

The Quantitative Effects of Socioeconomic Factors on Education

So far, the report has only presented the statistically significant factors individually. There are occasions, however, when more specific information is preferable. For example, the reader may want to know how much low family income and poor residential environment contribute to the poor educational attainment of individuals from poverty stricken neighbourhoods, after the effects of all other factors have been taken into account simultaneously. The Estimate Statement, which is one of the options of the ANOVA Generalized Linear Model,[20] deals directly with this type of question. The following discussion focuses on 18-year-olds from rural areas, but the results are generally applicable to other age groups and to urban areas.

In measuring the impact of favourable and unfavourable socioeconomic forces on the education of 18-year-olds, we use the terms "favourable" and "unfavourable" in a hypothetical, illustrative sense. An 18-year-old is said to come from a favourable socioeconomic background if he/she is from a family with an income of $75,000 or over, lives in a family-owned house valued at $150,000 or more, is a resident of Quebec, is from a family of multiple ethnic origins (e.g., a British father and a French mother), is a member of a four-person family, and lives in a husband and wife family. A youth is said to come from an unfavourable socioeconomic background if he/she is a member of a low income family (a family with total family income of less than $14,999); lives in a dwelling valued under $79,999; is a resident of Nova Scotia; is from a minority ethnic group; is from a family of six or more persons; and lives with relatives. Of course, in the real world, better or worse combinations are possible. These two extreme socioeconomic combinations are chosen strictly for illustrative purposes. They represent the ideal and worst possible worlds for individuals, limited by the sample size on the

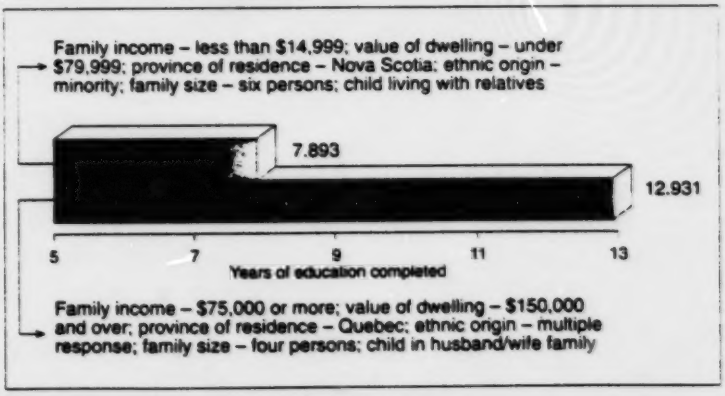
census tape. For example, it is possible to think of a better or worse combination of socioeconomic factors, but if there are no 18-year-olds with such socioeconomic attributes on the microdata tape, then the effect of this hypothetical socioeconomic background on education cannot be estimated.

Individuals from a favourable social and economic background may expect to complete 12.9 years of formal education by age 18 (Figure 6). The estimated figure is statistically significant at the level of 0.01 per cent; there is hardly any chance for the estimate to be a statistical artifact. At the other end of the spectrum, an individual from an unfavourable socioeconomic background can expect to complete only 7.9 years of formal education (Figure 6 and Table A-9). It should be mentioned that these estimates refer to average educational attainment. Certain individuals would of course accomplish more or less, depending on the personal traits. For example, if a person is more gifted than other members of the group, then the educational attainment of this individual will probably be higher than the estimated figure.

Figure 7 is a summary of the factorial contributions of the socioeconomic forces. The graph presents the net effect of each favourable factor in an additive, sequential fashion. The first part of the stack-bar graph illustrates that if the individual is from a family of unfavourable socioeconomic background, then the expected educational attainment for this 18-year old is 7.9 years of "education completed."

Figure 6

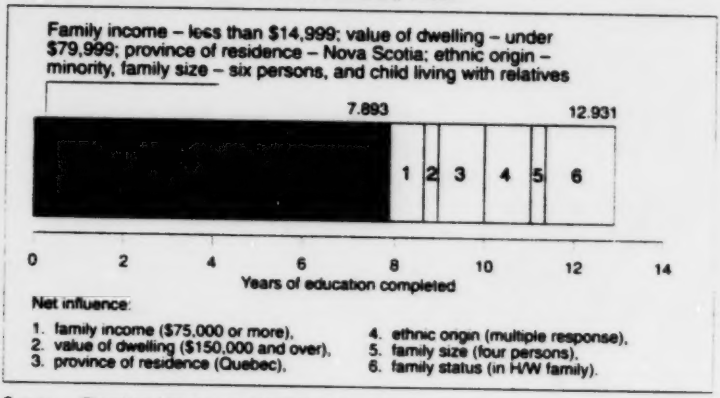
The effects of socioeconomic forces on the education of 18-year-olds from rural areas, Canada, 1986



SOURCE Based on data from the 1986 Population Census.

Figure 7

The effects of favourable socioeconomic factors on the education of 18-year-olds from rural areas, Canada, 1986



SOURCE Based on data from the 1986 Population Census.

However, if the individual were from a family with an annual income of $75,000 or more, with all other unfavourable factors unchanged, his/her educational attainment would be higher. The net increase in educational attainment is illustrated by rectangle 1 of the stack-bar graph. Similarly, rectangles 2 to 6 represent the net influence of other favourable factors on the individual's formal education. In the real world as well as in the statistical analysis, these factors affect education simultaneously. Therefore, the reader should take the graph simply as a presentation device. A husband-wife household has the largest impact on a child's education. Factors such as family income, province of residence, and ethnic origin also play a substantial role in determining a child's educational achievement. On the other hand, the contributions of family size and high property value are relatively modest.

In summary, the evaluation exercise has demonstrated that socioeconomic factors can make a great difference to a person's success in the educational system. Indeed, according to the example given here, the difference between very favourable and inordinately unfavourable socioeconomic factors can account for about five years of additional education completed. Clearly, if the country is interested in raising the educational standard of all Canadians, then, in addition to the perennial search for a more efficient education system, ways must be found to alleviate some of the unfavourable social and economic influence on less fortunate youths.

Concluding Remarks

This analysis has established that social and economic factors profoundly influence an individual's formal education. This finding is consistent with the consensus of past researchers in education economics. The contribution of this study is not that it has unearthed surprising results. On the contrary, its contribution is that it empirically confirms a popular belief. Furthermore, although past observers have acknowledged the influence of social and economic factors on education, they have seldom produced quantitative estimates for the contributions of a comprehensive set of individual socioeconomic factors. The present study may be the first research effort to do this.

From the list of significant variables found to affect educational attainment, there are few that the government has control over. For example, policymakers are in no position to change the economic status of poor families, to alter family size, or to provide single parent families with another parent.

A priori, poverty, large family size, and the lack of a husband/wife home are not directly responsible for low educational attainment. Only the overall environment created by these factors impede high educational attainment. How can the government help less fortunate youths to reach their potential? The U.S. Head Start programs have demonstrated that there are no "quick fixes." Preschool programs may help some underprivileged children for a short period. However, even the most enthusiastic supporters of Head Start do not expect the beneficial effects of the program to last throughout a child's academic career, for youths of low socioeconomic status are presented with too many chances of failing. Policymakers presently have no real answers to this problem. In North America, the children of wealthy families are often sent to private "prep schools" during their formative years. These fortunate youths tend to enjoy proportionately more success in the education system than their counterparts from the poor families. *A priori*, one must at least attribute part of their superior performance to the superior learning environments of "prep schools." Why, then, does the government not solve the socioeconomic problem by providing less fortunate children with a free "prep school" education? The answer, of course, is that the cost of such a program would be prohibitive. Given present fiscal stances, the federal and provincial governments simply cannot afford such a costly undertaking. Moreover, even if the government had the money to initiate and finance such a program, it is not clear that many poor parents would choose to send their children to free "prep schools," and they cannot be forced to do so. A free "prep school" program without participants can only become a white elephant. It would not solve the problem. If policymakers consider this solution impractical, then they should immediately consult the experts in education, psychology, criminology, and the police for a viable alternative.

Appendix

Summaries of Univariate Tests

Table A-1

Dependent variable — years of education completed; sample — all 18-year-olds from census metropolitan areas with populations of 100,000 or more, 1986

Independent variable (Source of variation)	F Value	Level of significance[1] (Per cent)
Income class	8.18	0.01
Value of dwelling	17.38	0.01
Province	3.22	0.67
Family size	6.95	0.01
Family status	4.01	0.13
Income class province	1.69	1.70
Income class family size	2.15	1.15

1 In Type III analysis, the main effect of an independent variable is adjusted for the effects of other independent variables and the interaction effects. For more details, see R.J. Freund, R.C. Littell, and P.C. Spector, *SAS System for Linear Models* (Cary: SAS Institute Inc., 1986), and *SAS/STAT User's Guide*, Version 6, Fourth Ed, (Cary: SAS Institute Inc., 1990).

SOURCE Based on data from the 1986 Population Census.

Table A-2

Dependent variable — years of education completed; sample — all 18-year-olds from small towns/cities and rural areas, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Income class	16.63	0.01
Value of dwelling	6.75	0.01
Province	14.78	0.01
Ethnic Origin	16.67	0.01
Family size	8.28	0.01
Family status	12.68	0.01

Table A-3

Dependent variable — years of education completed; sample — all 26-year-olds from census metropolitan areas with populations of 100,000 or more, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Income class	55.80	0.01
Value of dwelling	5.21	0.03
Ethnic origin	6.40	0.01
Family size	26.53	0.01
Province value of dwelling	1.71	1.68

Table A-4

Dependent variable — employment income in 1985; sample — all 26-year-olds from census metropolitan areas with populations of 100,000 or more, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Province	6.41	0.01
Family size	4.13	0.02
Education class	14.17	0.01
Weeks-worked class	556.77	0.01

Table A-5

Dependent variable — weeks worked in 1985; sample — all 26-year-olds from census metropolitan areas with populations of 100,000 or more, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Income class	133.57	0.01
Province	8.65	0.01
Family size	71.34	0.01
Education class	11.06	0.01
Income class education class	2.59	0.07

Table A- 6

Dependent variable — years of education completed; sample — all 26-year-olds from small towns/cities and rural areas, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Income class	25.98	0.01
Province	5.51	0.01
Ethnic Origin	10.15	0.01
Family size	44.42	0.01
Income class province	1.56	0.75

Table A-7

Dependent variable — employment income in 1985; sample — all 26-year-olds from small towns/cities and rural areas, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Province	6.25	0.01
Family size	5.59	0.02
Education class	16.06	0.01
Weeks-worked class	514.24	0.01

Table A-8

Dependent variable — weeks worked in 1985; sample — all 26-year-olds from small towns/cities and rural areas, Canada, 1986

Independent variable (Source of variation)	F Value	Level of significance (Per cent)
Income class	85.09	0.01
Province	7.55	0.01
Family size	73.20	0.01
Education class	15.66	0.01
Income class education class	1.98	1.34

Table A-9

Quantitative effects of socioeconomic factors on education, Canada, 1986

18-Year-olds from urban centres

Unfavourable influence

Family income — $14,999 or less in 1985
Value of dwelling — $79,999 or less
Family size — six or more persons per family
Living arrangement — child living with relatives

Estimated years of education completed — 9.610
(statistically significant at 0.01 per cent)

Favourable influence

Family income — $75,000 or more in 1985
Value of dwelling — $150,000 and over
Family size — four persons per family
Living arrangement — child living in husband/wife family
Estimated years of education completed — 12.422
(statistically significant at 0.01 per cent)

18-Year-olds from rural areas

Unfavourable influence

Family income — $14,999 or less in 1985
Value of dwelling — $79,999 or less
Province of residence — Nova Scotia
Ethnic origin — minority
Family size — six or more persons per family
Living arrangement — child living with relatives

Estimated years of education completed — 7.893
(statistically significant at 0.01 per cent)

Favourable influence

Family income — $75,000 or more in 1985
Value of dwelling — $150,000 and over
Province of residence — Quebec
Ethnic origin — multiple response (e.g., British and other)
Family size — four persons per family
Living arrangement — child living in husband/wife family

Estimated years of education completed — 12.931
(statistically significant at 0.01 per cent)

Notes

1 See James Coleman *et al.*, *Equality of Educational Opportunity* (Washington, D.C.: Department of Health, Education, and Welfare, 1966).

2 See James Coleman *et al.*, *Equality of Educational Opportunity*; and Burt S. Barnow and Glen G. Cain, "A reanalysis of the effect of head start on cognitive development: Methodology and empirical findings," *Journal of Human Resources* Spring 1977, 12(2): 177–97.

3 W. S. Barnett, "Benefit-cost analysis of the Perry Preschool Program and its long-term effects," *Educational Evaluation and Policy Analysis* 1985, 4(7): 333–42.

4 See Maisy Cheng *et.al.*, *The Every Secondary Student Survey* Fall 1987 (Toronto Board of Education, June 1989).

5 See Coleman *et al.*, *Equality of Educational Opportunity*.

6 See, for example, Barnow and Cain, "A reanalysis of the effect of Head Start."

7 Barnett, "Benefit-cost analysis of the Perry Preschool Program.

8 Richard J. Murnane, "What do effective teachers of inner-city children have in common?" *Social Science Research* March 1981, 10(1).

9 Anita A. Summers and Barbara L. Wolfe, "Do schools make a difference?" *The America Economic Review* September 1977, 67(4).

10 Richard J. Murnane *et. al.*, "Home resources and children's achievement" *Review of Economics* and *Statistics* August 1981, 63(3): 369–77.

11 See, for example, Summers and Wolfe, "Do schools make a difference?"

12 Maisy Cheng *et al.*, *The Every Secondary Student Survey*.

13 See, for example, Murnane *et al.*, "Home resources and children's achievement."

14 The census microdata file has separate data for 10 provinces and the Yukon/Northwest Territories (combined). Therefore, we only need 10 dummy variables to capture the provincial effects. The effect of the omitted territory is imbedded in the constant term of the estimated equation. 15ANOVA is a well developed theory in the history of statistics. The reader should have no problem locating some good books on the topic. See, for example, S.R. Searle *Linear Models* (New York: John Wiley & Sons, 1971), and F.A. Graybill, *Theory and Application of the Linear Model* (Belmont, Ca: Wadsworth Publishing, 1976).

16 This statistical property is not well known among the users of the multiple regression technique. For more information, see F.A. Graybill, *Theory and Application of the Linear Model*; R.J. Freund *et. al.*, *SAS System for Linear Models* (Cary: SAS Institute, 1986); and *SAS/STAT User's Guide*, Version 6, fourth ed. (Cary: SAS Institute, 1990).

17 In passing, we may point out that the Census information used in this study is always in the form of unbalanced design. The Generalized Linear Model of ANOVA enables us to deal with the data problem correctly and efficiently. See *SAS/STAT User's Guide*.

18 We did not ignore individuals aged 15–17 and 19–25 completely. For the sake of economizing the amount of statistical work, we divided young people into three age groups — 15–17, 19–22, and 23–25. This allowed us to study the behaviour of these individuals economically without introducing too many age-specific aggregation problems. The results for those aged 15–17 and 19–22 match fairly closely those for the 18-year-olds, and the results for the 23–25 year-olds are similar to those of the 26-year-olds. These results are available from the author upon request.

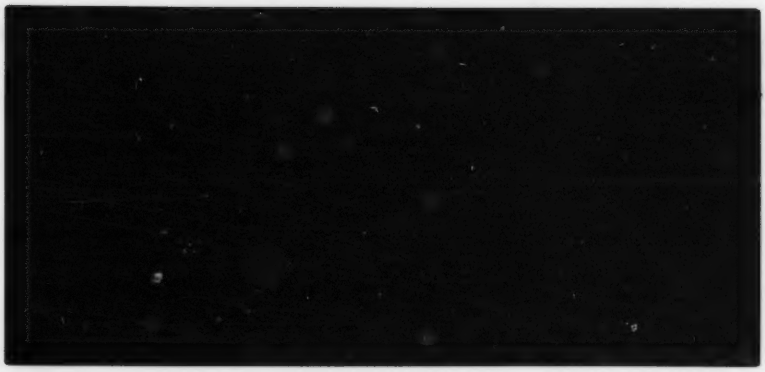
19 For more detail on this topic, see Freund *et. al.*, *SAS System for Linear Models*; and *SAS/STAT User's Guide*.

20 See *SAS/STAT User's Guide*.

Economic Council
of Canada

Conseil économique
du Canada

Working Paper
Document de travail



Document de travail n° 34

L'influence de la situation socio-économique sur le niveau d'instruction

Tom Siedule

1992

ISSN 1180-3495

L'influence de la situation socio-économique sur le niveau d'instruction

Les résultats publiés dans le présent document sont la seule responsabilité de l'auteur et, à ce titre, n'engagent en rien le Conseil économique du Canada.

Les documents de travail du Conseil économique sont protégés par le droit d'auteur de la Couronne et peuvent être cités dans les limites prescrites par la Loi canadienne sur le droit d'auteur. Toute demande d'autorisation d'en citer ou d'en reproduire de longs extraits doit être adressée au :

Directeur des publications
Conseil économique du Canada
Case postale 527
Ottawa (Ontario)
K1P 5V6

Table des matières

Remerciements vii

Avant-propos ix

Résumé xi

Introduction 1

Consensus et désaccord 1

Une question de terminologie 4

Perspective historique 5

Méthodologie 8

Conclusions empiriques 13
Les individus de 18 ans 14
Effets des facteurs individuels 14
Mise en garde 16
Les individus de 26 ans 17
Effets quantitatifs des facteurs socio-économiques sur le niveau d'instruction 18

Conclusion 21

Appendice 23

Notes 27

Liste des figures

1 Années de scolarité terminées par les individus de 18 ans, Canada, 1986 5
2 Années de scolarité terminées et revenu familial, individus de 18 ans, Canada, 1986 6

3 Années de scolarité terminées par les individus de 26 ans, Canada, 1986 7
4 Revenu du travail des individus de 26 ans, Canada, 1985 7
5 Semaines de travail des individus de 26 ans, Canada, 1985 8
6 Effets des forces socio-économiques sur le niveau d'instruction des individus de 18 ans des régions rurales, Canada, 1986 19
7 Effets des facteurs socio-économiques favorables sur le niveau d'instruction des individus de 18 ans des régions rurales, Canada, 1986 20

Remerciements

Je tiens à remercier Keith Newton, directeur de la recherche du projet de recherche du Conseil sur le système canadien d'éducation, d'avoir mis à ma disposition le temps et les moyens de développer la thématique ainsi que les travaux empiriques du présent projet. Ses commentaires, à différents stades de mon travail, en ont grandement amélioré la qualité. J'assume, il va sans dire, l'entière responsabilité des erreurs et des omissions.

Avant-propos

Depuis la publication du rapport Coleman, en 1966, aux États- Unis, décideurs, éducateurs et économistes s'accordent à reconnaître que les facteurs socio-économiques ont une grande influence sur le niveau d'instruction. Le présent rapport résume les conclusions d'un examen empirique de cette importante question. L'originalité de sa contribution est qu'il analyse un vaste échantillon de la population canadienne en employant comme variable dépendante le niveau d'instruction (années de scolarité). La plupart des études de ce genre, par contre, utilisent d'ordinaire comme variable dépendante des échantillons plus restreints pour évaluer les résultats de tests ou d'examens administrés à de jeunes enfants — plutôt que leur niveau d'instruction. Les estimations des effets sur l'instruction d'un ensemble exhaustif de facteurs socio-économiques individuels ne manqueront certes pas d'intéresser les décideurs. Il est à souhaiter que les résultats éclaireront également le public sur cette question.

Réalisée par Tom Siedule, économiste du Conseil, la présente étude s'inscrit dans le cadre du programme de recherche du Conseil sur le système canadien d'éducation.

Judith Maxwell
Présidente

Résumé

La présente étude vise principalement à déterminer si la situation socio-économique influe sur le niveau d'instruction d'un individu. Elle utilise à cette fin le fichier de micro-données du recensement de la population de 1986, de même que la technique de l'analyse de la variance. Les conclusions confirment non seulement la corrélation entre le niveau d'instruction et la situation socio-économique, mais elles mesurent aussi quantitativement l'effet d'une situation socio-économique inférieure sur le niveau d'instruction d'un individu. Selon cette étude, l'écart entre l'influence d'un milieu socio-économique très favorable et celle d'un milieu exceptionnellement défavorable peut être estimé à cinq années complètes de scolarité. Le message est des plus clairs : si le Canada souhaite rehausser le niveau d'instruction de tous les Canadiens, ses décideurs devront non seulement poursuivre leur recherche d'un système d'éducation plus efficace, mais aussi trouver des façons de contrecarrer certains des effets néfastes du milieu social et économique sur les jeunes moins privilégiés. Outre l'influence générale des facteurs socio-économiques sur le niveau d'instruction, la présente étude évalue leurs effets factoriels. Elle porte sur la population de toutes les provinces et des territoires. Ses conclusions devraient intéresser à la fois les éducateurs et les décideurs en matière de politique sociale.

Introduction

En 1966, Coleman *et al.* ont conclu que la situation socio-économique, et non l'école, exerçait l'influence la plus importante sur le niveau d'instruction[1]. La célèbre thèse de ce rapport voulant que « l'école ait peu d'influence sur l'instruction des enfants », demeure controversée, mais les chercheurs conviennent généralement que les facteurs socio-économiques exercent *a priori* une profonde influence sur le niveau d'instruction.

Les décideurs sont également conscients de l'effet de la situation socio-économique sur le niveau d'instruction des enfants. Financé par le gouvernement fédéral américain, le projet Head Start visait à relever le niveau d'instruction des enfants issus de milieux socio-économiques défavorisés. La question de savoir si ce programme a réussi à atteindre ses objectifs fait l'objet d'une controverse depuis sa mise en œuvre en 1965. Cependant, un certain nombre d'évaluations confirment que ce programme a eu un effet marqué sur les enfants des groupes minoritaires et sur les enfants blancs issus de familles monoparentales. Les résultats obtenus par les participants à ce programme étaient invariablement supérieurs à ceux des non-participants du même milieu socio-économique et des mêmes écoles[2]. De même, une analyse coûts-bénéfices du programme préscolaire Perry indique que la rentabilité de cet investissement dans le relèvement du niveau d'instruction des défavorisés était très élevée[3].

Les données canadiennes permettent de tirer une conclusion semblable[4]. Malheureusement, les études canadiennes portent principalement sur des conclusions empiriques ne touchant qu'un seul conseil scolaire ou une seule ville. Elles ne sont pas représentatives de l'ensemble du pays. Pour mieux connaître le niveau d'instruction de tous les jeunes Canadiens, il faut au préalable procéder à une étude nationale. En outre, la plupart des études existantes ne fournissent aux décideurs que la corrélation entre la situation socio-économique et le niveau d'instruction des enfants. Elles n'estiment pas de façon précise l'influence des facteurs socio-économiques individuels. Les travaux de recherche de l'étude contribueront peut-être à combler cette lacune.

Consensus et désaccord

Le projet Head Start constituait le programme gouvernemental le plus célèbre jamais mis en œuvre pour offrir aux enfants défavorisés d'âge préscolaire des mesures spéciales en milieu scolaire pour compenser les déficiences de leur milieu. Même si ce programme était également offert à des jeunes de situation socio-économique supérieure, le taux de parti-

cipation le plus élevé se retrouvait chez les individus dont la situation socio-économique était inférieure. Coleman *et al.* ont conclu que ce sont les enfants noirs issus de familles pauvres des États du Sud qui avaient tendance à bénéficier du programme Head Start[5]. Des recherches ultérieures ont montré que des jeunes défavorisés de race blanche et d'autres origines ethniques avaient également profité du programme Head Start[6]. Certains chercheurs soutiennent que les effets bénéfiques du programme étaient passagers et que les participants avaient tendance, après trois ans, à perdre les avantages acquis. Compte tenu du fait que les élèves ne participaient au projet Head Start que pendant un été ou un an, la question de savoir si ses effets étaient passagers ou permanents ne semble guère pertinente. Il s'agit de savoir si les enfants bénéficiaient de ce programme. Quand bien même ces bienfaits n'auraient duré qu'un an après la fin de la participation des enfants au programme, celui-ci n'en aurait pas moins prouvé son utilité. Pareille conclusion empirique donne tout simplement à penser qu'une situation socio-économique inférieure exerce une influence profondément nuisible sur le niveau d'instruction. Cela prouve également que les enfants de milieux socio-économiques défavorisés ont besoin d'être aidés davantage qu'une seule fois, au début de leurs études, pour exceller, ou simplement pour réaliser leur potentiel, au sein du système d'éducation.

Le programme préscolaire Perry, qui était un très petit projet semblable au projet Head Start, a également prouvé ses effets bénéfiques pour les élèves et la société. Une analyse coûts-bénéfices du programme établit le rapport coûts-bénéfices pour la société à six[7]. C'est-à-dire que pour chaque dollar investi dans ce programme, la société, à long terme, en retirerait six. Comment en arrive-t-on à ce résultat ? La réponse est simple : des recettes fiscales plus élevées et une réduction du coût des services sociaux suffisent amplement à compenser l'investissement. L'analyse coûts-bénéfices a suivi le développement des élèves participants et non participants depuis l'âge préscolaire jusqu'à 19 ans. Elle comprenait des données sur le niveau d'instruction atteint, l'emploi et les activités criminelles. Elle concluait qu'un plus grand nombre de participants étaient alphabétisés, avaient un emploi ou poursuivaient des études postsecondaires, et qu'un nombre plus faible de participants abandonnaient leurs études, étaient appréhendés pour délinquance ou à la suite d'actes criminels ou recevaient de l'aide sociale.

En un sens, le projet Head Start et le programme préscolaire Perry constituaient la réponse des décideurs aux problèmes socio-économiques des enfants défavorisés. En d'autres termes, les décideurs ont présupposé l'existence des problèmes et ont eu recours à ces programmes pour les résoudre. Ces programmes, par conséquent, n'apportaient pas une preuve directe de l'existence d'une influence socio-économique sur le niveau d'instruction.

Depuis la publication du rapport Coleman, les économistes ont tenté d'aborder directement la question de l'influence du milieu socio-économique. Tous s'efforcent, soit de prouver le bien-fondé de la thèse de Coleman, soit de la réfuter. La plupart des chercheurs soutiennent que les écoles exercent une influence très subtile. Par exemple, Murnane estime que les techniques pédagogiques et les caractéristiques générales des enseignants sont importantes lorsqu'il s'agit de prédire l'efficacité de l'enseignement[8]. Summers et Wolfe affirment que l'influence du milieu scolaire, des installations et des enseignants ne peut être établie qu'à partir des détails des micro-données[9]. Plus précisément, ils estiment que les élèves sous-performants, les élèves de familles à faible revenu et de race noire assimilent relativement bien certaines matières scolaires. Murnane *et al.*[10], en revanche, soulignent l'importance du milieu familial. Ils concluent que les façons différentes d'élever les enfants chez des mères de différents niveaux d'instruction ont une influence marquée sur les résultats scolaires des enfants, mais que le confort matériel à la maison ne semble pas influer de façon systématique sur leur instruction.

Bien que l'on semble s'accorder à reconnaître l'influence appréciable des facteurs socio-économiques sur le niveau d'instruction des individus, pareilles analyses présentent un certain nombre de problèmes méthodologiques. La plupart des études effectuées depuis le rapport Coleman mettent l'accent sur l'efficacité des écoles plutôt que sur l'efficacité des individus à l'intérieur du système d'éducation. Cette approche comporte certaines embûches. Par exemple, en règle générale, l'unité analytique doit être l'école; par conséquent, le niveau d'instruction des élèves est établi par les moyennes scolaires. Le caractère agrégatif de ces données tend à occulter la véritable influence de l'école sur les individus. Certaines études[10] reconnaissent l'insuffisance des « moyennes scolaires » et utilisent des micro-données portant sur les élèves[11], mais elles ne règlent pas un autre problème grave : leur champ d'observation géographique est limité. Du point de vue méthodologique, il est difficile de tirer une conclusion générale de ces études, car ce qui est vrai, par exemple, des districts scolaires de Philadelphie peut ne pas l'être d'autres districts scolaires. Certaines de ces études utilisent des unités de mesure brutes de l'influence socio-économique (p. ex. elles utilisent les termes « noirs » ou « blancs » pour obtenir une approximation des différences de races, et le revenu de la famille de l'élève pour décrire toutes les influences socio-économiques) et négligent trop d'éléments importants de la situation socio-économique.

Le présent rapport ne prétend pas réfuter les études existantes. Il tente plutôt de donner au lecteur un aperçu du débat en cours. Il existe certaines lacunes. Par exemple, une étude à l'échelle nationale de l'influence des facteurs socio-économiques sur le niveau d'instruction aurait dû être effectuée depuis longtemps; en outre, si l'argument des utilisateurs de

micro-données est valide, seule une analyse d'une base de micro-données à l'échelle nationale peut nous fournir un tableau complet. La présente étude examine cet aspect relativement inconnu de l'économie de l'éducation.

Idéalement, une étude du rendement du système d'éducation devrait englober l'influence des facteurs institutionnels aussi bien que personnels. Malheureusement, les organismes de statistique n'ont pas réussi à intégrer ces deux dimensions dans une base de micro-données nationales. La base de données existante au Canada comprend des renseignements sur le niveau d'instruction des individus et leurs caractéristiques socio-économiques. Ces données, cependant, ne font pas partie des données scolaires. À l'heure actuelle, il n'existe pas de méthode acceptable permettant aux chercheurs de relier l'information sur les individus aux données scolaires.

La présente étude utilise les micro-données disponibles du recensement de la population de 1986 pour étudier les taux de réussite et d'échec des Canadiens au sein du système d'éducation. En raison des limites des données, l'étude ne porte que sur la dimension quantitative de l'influence socio-économique.

Une question de terminologie

En économie de l'éducation, de nombreux chercheurs emploient la même expression pour décrire des concepts fort différents. Par exemple, l'expression « rendement scolaire » peut avoir des significations différentes selon les auteurs. Certains l'emploient pour désigner la moyenne pondérée cumulative d'un élève; d'autres s'en servent pour parler de la moyenne pondérée cumulative d'une classe, d'une école ou d'un conseil scolaire. Un autre exemple est l'expression « rendement du système d'éducation ». Cela peut signifier le nombre d'élèves excellents, bons, moyens ou médiocres produits par un établissement d'enseignement. Toutefois, certains analystes utilisent cette expression pour décrire les variations entre les présences à l'école, les comportements ou les taux de délinquance des élèves.

Nous ne nions pas l'importance des multiples aspects de l'éducation. Dans la présente étude, cependant, le caractère limité des données nous oblige à employer l'expression « niveau d'instruction » au sens de nombre d'années de scolarité terminées par un individu et l'expression « facteurs socio-économiques » au sens de facteurs sociaux, économiques, démographiques et géographiques.

Perspective historique

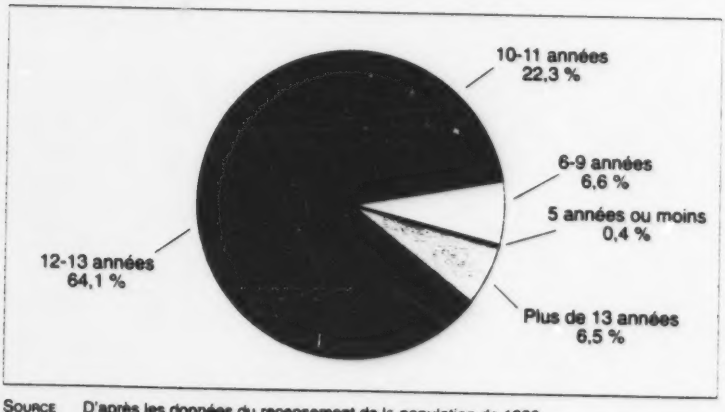
En 1986, plus de 29 % des Canadiens âgés de 18 ans avaient terminé un maximum de 11 années de scolarité, tandis que 64 % avaient terminé de 12 à 13 années, et plus de 6 %, plus de 13 années (figure 1). Les « années de scolarité terminées » désignent le niveau d'instruction (classes et années terminées) dans les écoles élémentaires et secondaires, plus le total des années de scolarité terminées dans des établissements d'enseignement post-secondaires. Si un individu a mis 14 ans à terminer la douzième année, on ne lui comptera que 12 années de scolarité. En revanche, un élève qui terminerait sa troisième année en 11 ans se verrait reconnaître 13 années de scolarité.

Les statistiques révèlent l'existence d'un certain nombre de « performants » et de nombreux « sous-performants ». Au Canada, les enfants entrent d'ordinaire en première année à l'âge de six ans. Par conséquent, à l'âge de 18 ans, ils devraient avoir terminé plus de 11 années de scolarité. La question qui se pose est de savoir pourquoi le tiers des individus âgés de 18 ans n'avaient pas réalisé leur potentiel au sein du système d'éducation en 1986.

Un examen rapide des données disponibles révélerait que la situation économique (reflétée par le revenu familial) a eu un effet sur le niveau d'instruction moyen des individus de 18 ans (figure 2). Les jeunes issus de familles à revenu élevé comptaient, en moyenne, plus d'années de scolarité que leurs homologues provenant de familles à revenu faible.

Figure 1

Années de scolarité terminées par les individus de 18 ans, Canada, 1986



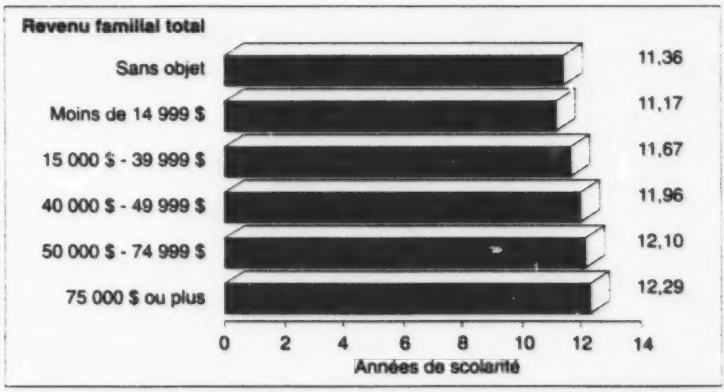
Source D'après les données du recensement de la population de 1986.

D'un point de vue méthodologique, cependant, il y a lieu de s'interroger sur les implications de ces données brutes. Tout d'abord, l'écart est-il assez grand pour être statistiquement significatif ? Deuxièmement, le revenu familial n'est qu'un des nombreux facteurs qui peuvent influer sur le niveau d'instruction. Ainsi, l'origine ethnique, le fait de vivre en milieu urbain ou rural, le milieu résidentiel, la province de résidence, la taille de la famille, et le fait que celle-ci comprenne un ou deux conjoints peuvent également avoir un effet important sur le niveau d'instruction d'un individu. Les chercheurs doivent tester systématiquement ces hypothèses avec les données et les méthodes statistiques appropriées.

En matière de niveau d'instruction, les individus âgés de 26 ans étaient comparables à ceux de 18 ans — 27 % d'entre eux comptaient 11 années ou moins de scolarité et 9 % n'étaient pas parvenus à la dixième année (figure 3). Dans le cas des individus de 26 ans, la mesure de la réussite ne doit pas se cantonner au système d'éducation, mais inclure l'expérience acquise sur le marché du travail. La majorité des individus de ce groupe d'âge auraient dû avoir terminé leurs études en 1985. Leur revenu du travail et le nombre de semaines de travail au cours d'une année devraient par conséquent être de bons indices de leur rendement sur le marché du travail. Apparemment, bon nombre des individus âgés de 25 ans en 1985 ne gagnaient pas bien leur vie cette année-là. Au total, 46 % ont gagné moins de 10 000 $ en 1985; 33 % ont touché un revenu annuel de 5 000 $ ou moins (figure 4). Les statistiques sur le nombre de semaines de travail recoupent les données concernant les revenus du travail. En effet, 35 % des membres de ce groupe d'âge ont travaillé

Figure 2

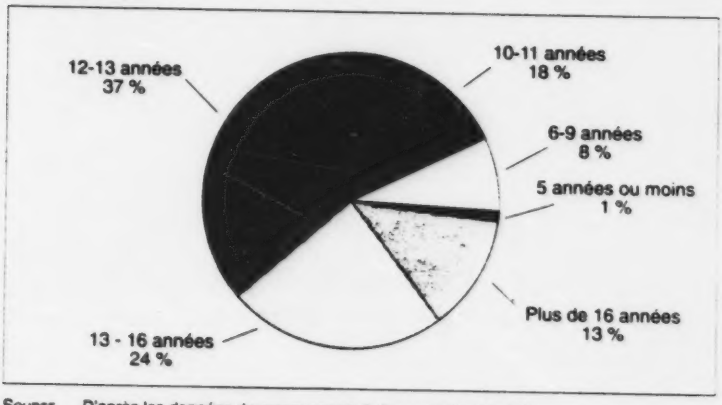
Années de scolarité terminées et revenu familial, individus de 18 ans, Canada, 1986



Source D'après les données du recensement de la population de 1986.

Figure 3

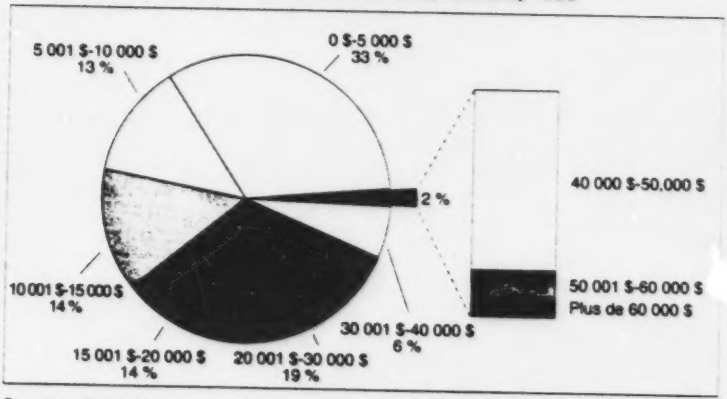
Années de scolarité terminées par les individus de 26 ans, Canada, 1986



SOURCE D'après les données du recensement de la population de 1986.

Figure 4

Revenu du travail des individus de 26 ans, Canada, 1985

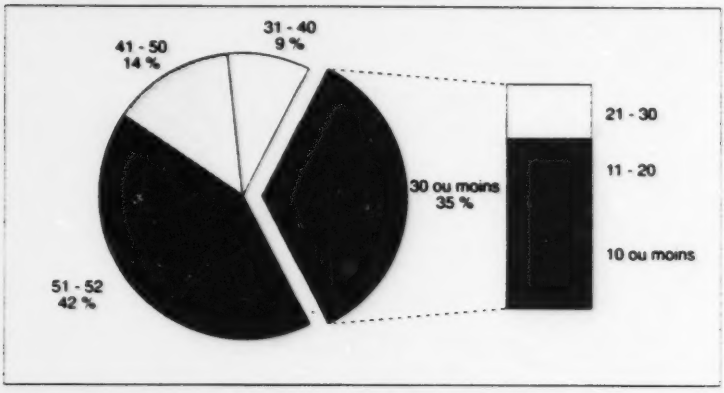


SOURCE D'après les données du recensement de la population de 1986.

moins de 30 semaines en 1985, y compris une majorité qui a occupé un emploi rémunéré pendant moins de dix semaines (figure 5). Un niveau d'instruction inférieur, un faible revenu et un chômage involontaire sont-ils des problèmes liés ? Peuvent-ils être imputés aux mêmes facteurs socio-économiques ? Telles sont certaines des questions auxquelles le présent document espère répondre.

Figure 5

Semaines de travail des individus de 26 ans, Canada, 1985



SOURCE D'après les données du recensement de la population de 1986.

Méthodologie

Tout chercheur effectuant une étude empirique doit se servir du contenu d'une base de données. Comme la plupart des données de la présente étude proviennent du recensement de la population de 1986, les travaux empiriques doivent être définis en fonction de ce qui se trouve dans le recensement. Pour cette raison, le nombre d'années de scolarité terminées a été retenu comme variable dépendante pour les individus de 15 à 22 ans. Pour les individus âgés de 23 à 26 ans, trois variables dépendantes sont prises en compte — les années de scolarité terminées, le revenu du travail gagné en un an et le nombre de semaines de travail au cours d'une année — afin de déterminer ce que peut signifier la « réussite » pour des individus plus âgés, à ce stade de leur carrière. Le recensement contient suffisamment de renseignements pour permettre aux chercheurs de calculer chacune de ces trois variables dépendantes. En ce qui concerne les années de scolarité, le recensement contient des renseignements sur le grade, certificat ou diplôme le plus élevé obtenu par un individu; la dernière année terminée au niveau élémentaire ou secondaire; le total des années terminées dans un établissement autorisé à décerner des grades et dans un établissement non autorisé. Le recensement donne également des renseignements sur les semaines de travail et le revenu du travail gagné en 1985.

Le recensement fournit aussi des données sur un certain nombre de variables explicatives (ou indépendantes) possibles, telles que l'âge, le sexe, la province de résidence, le fait de vivre en milieu urbain ou rural,

la situation de famille de recensement ou structure familiale, l'origine ethnique, le fait que l'individu ait travaillé à temps plein ou à temps partiel en 1985, le nombre d'heures de travail pendant la semaine de référence, l'activité sur le marché du travail, les principales matières étudiées, le nombre de membres dans une famille de recensement, le revenu total de la famille de recensement et la valeur du logement. Il est évident qu'un grand nombre de ces variables sont des variables de classification. Par exemple, la « province de résidence » n'est pas, par définition, une donnée mesurable et doit être codée sous un nom de classification ou comme nombre discret. Même le revenu familial total, dans les micro-données du recensement, est codé par tranches de revenu, et non pas selon le montant exact qu'a gagné cette famille en 1985. Heureusement, les variables indépendantes conviennent parfaitement à l'analyse statistique de la variance.

Nous posons l'hypothèse que le niveau d'instruction d'un individu dépend de six facteurs socio-économiques : le revenu familial, la valeur marchande du logement, l'origine ethnique, la province de résidence, la taille de la famille, ainsi que la situation de famille ou structure familiale. Le revenu familial constitue un facteur socio-économique évident; l'influence des autres variables indépendantes, en revanche, nécessite quelques explications. La valeur marchande du logement où habite l'enfant est une approximation des effets du milieu où sont situés le logement et l'école. Exception faite des enfants de familles riches qui sont envoyés dans des écoles privées, les jeunes Canadiens fréquentent généralement l'école de leur quartier. Évidemment, un quartier aisé va généralement de pair avec un milieu plus propice à l'effort intellectuel, tandis qu'une école située dans un « ghetto » risque de connaître plus de problèmes de toxicomanie et de délinquance. L'influence de l'origine ethnique sur le niveau d'instruction des enfants a été empiriquement établie par Cheng *et al.*[12]. Les effets de la province de résidence sont moins évidents. Ils peuvent se manifester de deux façons. Tout d'abord, les différences de culture et de conditions entre les provinces influent sur la perception qu'a un individu de la scolarité optimale nécessaire pour son travail (p. ex. un agriculteur des Prairies n'estimerait probablement pas essentiel d'être titulaire d'un doctorat pour élever du bétail). Ensuite, les provinces ne consacrent pas les mêmes ressources à l'éducation (les paiements de péréquation du gouvernement fédéral au titre de l'éducation visent à atténuer ces différences, mais ne parviennent pas vraiment à placer toutes les provinces sur un pied d'égalité). Ces écarts entre les provinces devraient donc se répercuter sur le niveau d'instruction de leur population. La « taille de la famille » influe sur le niveau d'instruction des enfants d'une façon particulière[13]. Toutes choses étant égales par ailleurs, nous tenons pour acquis qu'il existe une taille optimale de la famille qui favorise la réussite scolaire. Un enfant unique n'a pas de frère ou sœur avec qui partager ses expériences scolaires, ce qui risque de nuire

à son épanouissement. À l'opposé, si une famille compte trop d'enfants, ceux-ci doivent se partager non seulement des ressources financières limitées, mais aussi le peu de temps dont disposent leurs parents pour les guider. Enfin, la variable concernant la structure familiale (si l'enfant vit dans une famille comptant deux conjoints, une famille monoparentale, avec des parents ou des personnes sans lien de parenté) se passe d'explication.

Bien que les effets quantitatifs de ces variables doivent être vérifiés de façon empirique, peu de chercheurs en dénieraient la valeur explicative. Il existe cependant d'autres variables socio-économiques qui, en théorie, se prêtent bien à notre hypothèse. La liste présentée ici représente les variables qui interviennent dans la version finale de l'enquête statistique. L'étude exclut les autres variables pour diverses raisons : 1) elles ne figurent pas dans l'échantillon de 2 % du recensement de 1986; 2) ces variables avaient été retenues au départ, mais ont été abandonnées par la suite, parce qu'elles n'avaient pas d'influence statistiquement significative sur le niveau d'instruction d'un individu; ou 3) elles n'étaient pas essentiellement de nature socio-économique.

En théorie, rien n'empêche le chercheur d'appliquer à l'analyse ci-haut la méthode de la régression multiple. En pratique, cependant, cette approche présente plusieurs difficultés. Par exemple, le chercheur peut faire en sorte que la variable dépendante soit égale au nombre d'« années de scolarité terminées » et que les variables indépendantes correspondent aux caractéristiques socio-économiques d'un individu, lesquelles peuvent être exprimées sous la forme de grands ensembles de variables auxiliaires en fonction de « 0 » et de « 1 ». La variable dépendante ne présente aucun problème, étant donné que le niveau d'instruction de l'enfant est mesuré en fonction du nombre d'années de scolarité terminées. Les variables indépendantes sont cependant plus compliquées. Dans le fichier de micro-données du recensement, le revenu familial se présente sous la forme de 24 tranches de revenu. La variable touchant la province de résidence est semblable. L'enfant doit résider dans l'une des dix provinces, dans les Territoires du Nord-Ouest ou au Yukon. Par conséquent, pour la seule variable de la province de résidence, le chercheur a besoin de dix variables auxiliaires pour obtenir les effets de la province sur le niveau d'instruction[14]. En plus de cinq ensembles de variables auxiliaires représentant cinq caractéristiques socio-économiques, le chercheur doit tenir compte de leur interaction. Par exemple, le revenu familial moyen est beaucoup plus élevé en Ontario qu'en Saskatchewan. Par conséquent, on peut s'attendre qu'un revenu familial de 50 000 $ par an n'exercera pas la même influence sur le niveau d'instruction des individus de ces deux provinces, ce qui constitue, techniquement, l'interaction des variables revenu familial et province de résidence. De même, le chercheur peut étudier les effets de l'interaction d'autres caractéristiques socio-

économiques. L'équation d'estimation finale pourrait facilement comprendre plus de 150 variables auxiliaires et interactions. En économétrie appliquée, une équation comportant autant de variables explicatives (ou indépendantes) est généralement considérée comme peu pratique. D'autres raisons, cependant, s'opposent à l'application de la méthode de la régression multiple à ce cas précis. Par exemple, si l'influence de la province de résidence sur le niveau d'instruction d'un enfant n'est pas différente de la moyenne nationale, le coefficient estimé pour cette variable auxiliaire devient statistiquement non significatif. En revanche, l'influence de certaines provinces pourrait être sensiblement différente de la moyenne nationale, et les coefficients estimés pour ces variables auxiliaires peuvent alors être très significatifs. Ainsi, les coefficients estimés nous fournissent certaines indications quant aux effets des provinces individuellement. Cependant, nous n'avons toujours pas de réponse claire à la question : « La province de résidence a-t-elle un effet sur le niveau d'instruction d'un enfant ? » Le chercheur ne peut répondre à cette question sans effectuer un nouveau test de la signification statistique de *l'ensemble* des coefficients des variables auxiliaires des provinces (c.-à-d. administrer son test après estimation de l'équation). En outre, le chercheur voudra parfois savoir si l'effet d'une région (p. ex. la région de l'Atlantique) est sensiblement différent ou non de celui d'une autre région (p. ex. l'Ontario). L'équation d'estimation ne lui fournira pas cette réponse sans de plus amples calculs statistiques. Ainsi, en pratique, la « camisole de force » de la régression multiple, appliquée à l'hypothèse socio-économique, ne constitue pas la meilleure méthode statistique.

L'analyse de la variance est une méthode statistique qui peut servir à étudier les effets de variables indépendantes multiples (toutes sous forme de classification) sur une ou plusieurs variables dépendantes (toutes sous forme mesurable). Lorsque le chercheur relie une variable dépendante à un ensemble de variables indépendantes, il applique l'analyse à une variable de la variance. Lorsqu'il tente de mesurer les effets de variables indépendantes multiples sur une ou plusieurs variables dépendantes simultanément, il applique l'analyse à plusieurs variables de la variance. Aux fins de la présente étude, nous avons appliqué aux données concernant les individus âgés de 15 à 22 ans l'analyse à une variable de la variance. Aux données sur les individus âgés de 23 à 26 ans, nous avons appliqué à la fois l'analyse à une variable et l'analyse à plusieurs variables, parce que notre étude porte non seulement sur le niveau d'instruction de ce groupe d'âge, mais également sur le revenu du travail et les semaines de travail. Les résultats empiriques de ces deux méthodes d'analyse se recoupent, mais nous avons constaté que les conclusions de l'analyse à plusieurs variables ne nous éclairaient pas davantage sur la question. C'est pourquoi le présent rapport ne fera plus état des résultats obtenus par l'analyse à plusieurs variables. Pour la suite du rapport, l'expression analyse de la variance sera employée strictement au sens d'analyse à une variable[15].

L'hypothèse socio-économique énoncée ci-haut est typique du genre de problème que les chercheurs soumettent à l'analyse de la variance. Techniquement, l'hypothèse nulle veut que les facteurs socio-économiques n'entraînent aucun effet sur le niveau d'instruction des individus. En conséquence, la proposition opposée voudrait que ces facteurs exercent une influence. Nous nous proposons d'inclure dans le test six facteurs socio-économiques. Chaque facteur (ou variable indépendante) comporte plusieurs classes (p. ex. la variable province de résidence en compte 11), et chaque facteur peut interagir avec les autres. La méthode statistique employée, l'analyse à degré élevé de la structure de la variance, est en mesure de tenir compte simultanément des effets de tous les facteurs socio-économiques.

Dans une analyse statistique, outre le fait que toutes les variables dépendantes doivent être sous une forme mesurable et toutes les variables indépendantes sous forme de classification, le chercheur doit connaître les caractéristiques des données : c'est-à-dire, sont-elles équilibrées ou non équilibrées ? La méthode de l'analyse de la variance dans le cas de *données équilibrées* est fort différente de celle employée pour des données non équilibrées[16]. Appliquer la mauvaise méthode à des données non équilibrées peut donner lieu à des conclusions trompeuses. Malheureusement, en matière de données économiques, la forme non équilibrée est plutôt la règle que l'exception. Par exemple, si le chercheur pose l'hypothèse que les jeunes issus de familles à revenu moyen ou élevé ont tendance à obtenir un niveau d'instruction plus élevé que les jeunes provenant de familles à faible revenu, il s'agit d'un problème simple d'analyse de la variance. Toutefois, cela ne règle pas la question des données équilibrées et non équilibrées. Le chercheur devra se renseigner à ce sujet en comptant le nombre d'individus à l'intérieur de chaque case du plan d'analyse de la variance. Si le nombre de jeunes provenant de familles à faible revenu n'est pas égal au nombre de jeunes issus de familles à revenu moyen ou élevé, le problème devient automatiquement une analyse simple de la variance à données non équilibrées[17].

Le modèle linéaire généralisé d'analyse de la variance nous permet d'estimer le niveau d'instruction attendu d'un individu, sous réserve de l'influence d'une combinaison précise de facteurs socio-économiques (p. ex. une combinaison de multiples facteurs défavorables ou une combinaison des facteurs les plus favorables). Ce modèle peut être employé pour la réalisation statistique d'une analyse de la variance. Non seulement fournit-il des données sur le nombre estimé d'années de scolarité terminées par les individus en question, mais il nous indique également la mesure dans laquelle cette estimation est statistiquement significative. En bref, aux fins de la présente étude empirique, l'analyse de la variance constitue une méthode plus pratique et plus productive que la traditionnelle régression multiple, car elle permet d'analyser des variables chif-

frées (ou dépendantes) et des variables à classifications multiples (ou indépendantes) et produit une série de statistiques utiles et d'équations d'estimation de coefficients (les équations d'estimation sont fournies par le modèle linéaire généralisé d'analyse de la variance).

Pour faciliter la description, le présent rapport décrit et discute principalement les données historiques et les résultats statistiques concernant les individus âgés de 18 et de 26 ans. Il comprend des données de toutes les provinces et des territoires. En raison du champ d'observation de cette étude et de la méthode d'application de l'analyse de la variance aux microdonnées, elle constitue peut-être la première du genre sur l'instruction au Canada.

Conclusions empiriques

A priori, les jeunes des centres urbains ont des aspirations différentes de ceux des milieux ruraux en matière de scolarité. Pour cette raison, l'analyse empirique classifie tous les individus âgés de 15 à 26 ans selon qu'ils habitent en milieu urbain ou rural. En d'autres termes, nous analysons séparément les sous-populations urbaines et rurales. Ces deux sous-populations sont : a) urbaine — soit les individus qui habitaient des régions métropolitaines de recensement ayant un noyau urbain de 100 000 habitants ou plus en 1986 et b) rurale — tous les autres individus qui habitaient dans des villages, des villes et des régions rurales. Cette stratification sépare automatiquement l'effet du milieu urbain de celui du milieu rural sur le niveau d'instruction. Les exercices d'analyse de la variance ont dès lors pour but de déterminer l'influence des autres facteurs sur la scolarité.

La présente étude tient compte du rapport entre la scolarité et l'âge. L'analyse statistique porte séparément sur les données concernant les individus âgés de 18 ans et ceux âgés de 26 ans. Ces âges représentent deux stades critiques du développement en matière de scolarité et de carrière. À l'âge de 18 ans, un individu doit d'ordinaire décider d'entrer pour la première fois sur le marché du travail ou de poursuivre ses études et sa formation. À l'âge de 26 ans, une personne a généralement terminé ses études et peut même faire partie de la population active depuis quelques années déjà. Les analyses statistiques de ces deux groupes d'âge sont par conséquent essentielles à notre thèse socio-économique. Les résultats de ces analyses statistiques concernant les individus âgés de 18 et de 26 ans figurent par conséquent au tableau A-1 à A-8[18].

Les individus de 18 ans

Comme le montrent les tableaux A-1 et A-2, les facteurs utilisés pour expliquer le niveau d'instruction des individus de 18 ans des milieux urbain et rural sont très significatifs d'après les critères statistiques habituels. Les statistiques sont soumises au modèle linéaire généralisé de type III d'analyse de la variance. Dans ce type d'analyse, l'effet principal d'un facteur (ou variable indépendante) est corrigé en fonction des effets d'autres facteurs et des effets d'interaction. L'analyse statistique nous fournit également les résultats d'autres types d'analyse de la variance qui ne tiennent pas complètement compte des effets corrigés utilisés pour l'analyse de type III[9]. Ces résultats ne sont pas consignés ici. Qu'il suffise de souligner que les résultats statistiques obtenus par les autres types d'analyse sont au moins aussi prononcés que ceux qui figurent aux tableaux A-1 et A-2. Comme l'indiquent ces tableaux, outre les différences évidentes de signification statistique, l'origine ethnique semble exercer une influence sensible sur le niveau d'instruction des jeunes de 18 ans des régions rurales, mais non sur celui des jeunes des grands centres urbains. Il se peut que le creuset des grandes villes assimile les jeunes de différentes origines ethniques et atténue l'effet du milieu culturel. Le corollaire de cette proposition est que les jeunes d'origines ethniques différentes, issus de régions rurales n'ont pas suffisamment l'occasion de rencontrer des gens d'autres milieux culturels. C'est pourquoi ils ont tendance à conserver leur culture d'origine. Les interactions de la classe de revenu avec la province de résidence et de la classe de revenu avec la taille de la famille se révèlent statistiquement significatives pour les jeunes des grands centres urbains, mais non significatives pour les individus de 18 ans habitant dans les régions rurales. Étant donné que la classe de revenu, la province de résidence et la taille de la famille exercent en soi une influence appréciable sur le niveau d'instruction des jeunes de 18 ans des milieux urbain et rural, les interactions traduisent les effets en deça et au-delà de ceux des trois variables séparément. Les effets de l'interaction illustrent théoriquement l'influence bénéfique ou nuisible de la combinaison des caractéristiques socio-économiques.

Effets des facteurs individuels

Revenu familial — Parmi les jeunes de 18 ans des grandes régions métropolitaines de recensement et des régions rurales (à l'exception de la classe « sans objet » qui regroupait des jeunes vivant en communauté, dans des familles à l'étranger, les résidants temporaires et les personnes ne faisant pas partie de la famille), ceux qui étaient issus de familles aisées totalisaient plus d'années de scolarité que ceux qui provenaient de familles pauvres. Le niveau d'instruction des jeunes issus de familles dont le revenu s'établissait entre 40 000 $ et 49 999 $ ou plus est statistiquement

plus élevé que celui des individus dont les familles avaient un revenu inférieur à 14 999 $. L'influence du revenu familial sur le niveau d'instruction est ainsi démontrée.

Valeur du logement — Cette variable est fort significative, mais ses effets sur le niveau d'instruction doivent être interprétés avec prudence. Les jeunes dont les parents sont locataires tendent à avoir moins d'instruction que ceux dont les parents sont propriétaires. Cependant, le niveau d'instruction des jeunes vivant dans des logements coûteux n'est qu'à peine supérieur à celui des jeunes habitant des demeures moins chères. Le test de contraste de l'analyse de la variance confirme que la différence entre un logement en propre et un logement loué entraîne la signification statistique de cette variable. Par conséquent, tant pour les milieux urbain que rural, cette variable traduit les différents effets d'un logement en propre par rapport à un logement loué, plutôt que l'influence de la valeur du logement sur la scolarité.

Province de résidence — La variable province de résidence influe sensiblement sur le niveau d'instruction des jeunes des grands centres urbains et des régions rurales, mais la signification statistique est plus faible pour le milieu urbain que pour le milieu rural. Les jeunes des grands centres urbains du Québec et de l'Ontario réussissent un peu mieux que leurs homologues du Manitoba, de la Saskatchewan, de l'Alberta et de la Colombie-Britannique. (Il convient de noter que l'échantillon de 2 % du recensement de 1986 ne comprend aucune région métropolitaine de recensement située dans les provinces de l'Atlantique, au Yukon ou dans les Territoires du Nord-Ouest. Par conséquent, l'échantillon de la population urbaine ne couvre pas les individus de ces provinces et territoires.) Le niveau d'instruction des jeunes des régions rurales variait davantage d'une province à l'autre. En général, les individus de 18 ans des provinces de l'Atlantique comptaient moins d'années de scolarité que ceux des autres provinces. Les jeunes du Yukon et des Territoires du Nord-Ouest ont tendance à avoir la plus faible scolarité.

Origine ethnique — Le recensement de 1986 répartit les individus entre 28 groupes ethniques. Cependant, si les données sont classées par recoupement selon qu'il s'agit d'un milieu urbain ou rural, de nombreuses cases ne contiennent guère de renseignements. Pour cette raison, les 28 groupes ethniques ont été regroupés en cinq grands groupes. Bien que ce regroupement contourne le problème des données, il en crée un autre dans le groupe des « minorités » qui consiste alors en individus de toutes sortes d'origines ethniques, y compris des Asiatiques de l'Ouest, des Arabes, des Asiatiques du Sud, des Chinois, des Philippins, des Asiatiques de l'Est et du Sud-est, des autochtones, des Noirs, des Noirs d'Afrique et des Antilles. En raison de cette agrégation, l'analyse statistique ne

permet pas de saisir la diversité des groupes d'origine ethnique minoritaire. À cause de cette limite des données, la variable origine ethnique n'est significative que pour les régions rurales.

Taille de la famille — Tant dans les milieux urbains que ruraux, les jeunes issus de familles de quatre personnes ont tendance à compter plus d'années de scolarité que ceux provenant de familles plus nombreuses ou moins nombreuses. Abstraction faite de la taille de la famille, les jeunes qui vivent avec leur famille ont tendance à réussir mieux que ceux du même âge qui vivent hors-famille. Dans l'ensemble, la conclusion est que les familles de quatre personnes sont les plus propices à la poursuite des études des enfants alors que les familles nombreuses ou très petites, ou un milieu hors-famille semblent moins propices à la poursuite des études.

Situation familiale et structure familiale — Les enfants issus de familles comptant deux conjoints ou de familles monoparentales dont le chef est le père ont plus de chance de poursuivre leurs études que ceux qui vivent avec une mère seule, des membres de leur famille, des personnes étrangères à leur famille, des conjoints de fait ou seuls. Tout cela souligne l'importance du milieu de vie pour le niveau d'instruction, et ce tant en milieu urbain qu'en milieu rural.

Mise en garde

Les liens établis entre le niveau d'instruction et les facteurs économiques impliquent des rapports extrêmement complexes. Les facteurs socio-économiques peuvent influencer, mais ne peuvent pas entraîner la réussite ou l'échec d'un individu au sein du système d'éducation. En réalité, le revenu familial reflète en gros plusieurs facteurs. Dans certains cas, il traduit le niveau d'instruction des parents, l'aptitude des parents à organiser et employer leur temps convenablement et leur capacité d'orienter leurs enfants dans la « jungle » du système d'éducation. Si tel est le cas, c'est la valeur implicite du revenu familial et non la valeur monétaire qui influe sur le niveau d'instruction de l'enfant. De même, c'est le milieu créé par la valeur du logement, la taille de la famille, la situation familiale ou structure familiale qui contribuent à la réussite ou au succès d'un individu au sein du système. Il résulte de tout cela que les décideurs ont peu de choix pratiques lorsqu'il s'agit de remédier aux effets d'une situation socio-économique défavorable. Un niveau d'instruction élevé nécessite des installations éducatives adéquates et des milieux favorables pour tous les jeunes Canadiens. Jusqu'ici, ni l'État ni les éducateurs n'ont été en mesure de mettre ces avantages à la portée des jeunes moins bien nantis.

Les individus de 26 ans

Comme il a été mentionné plus haut, dans le cas des individus de 26 ans, l'étude tente de déterminer le succès dans la vie, soit la réussite au sein du système d'éducation et sur le marché du travail. Par conséquent, la recherche consiste à faire ressortir les facteurs socio-économiques qui peuvent influer de façon appréciable sur le nombre d'années de scolarité terminées, le revenu du travail et le nombre de semaines de travail au cours d'une année. Étant donné que le niveau d'instruction peut influer sur le revenu et l'emploi d'un individu et que le nombre de semaines de travail se répercute directement sur le revenu du travail, il faut établir de nouvelles variables de classification lorsqu'elles sont employées comme variables explicatives. La méthode de recherche doit notamment regrouper les années de scolarité terminées selon trois classes. Les individus comptant moins de 10 années de scolarité constituent la classe 1, ceux qui comptent plus de 11 mais moins de 12 années de scolarité forment la classe 2 et ceux qui comptent 13 années ou plus de scolarité constituent la classe 3. Les semaines de travail sont codées de la même façon, s'il y a lieu, sous la forme d'une variable de classification. Les tableaux A-3 à A-8 résument les résultats concernant les individus de 26 ans des milieux urbain et rural et chacune des variables dépendantes séparément.

Les résultats de la variable concernant le niveau d'instruction sont compatibles avec ceux des individus de 18 ans. Le revenu familial, la valeur du logement, l'origine ethnique et la taille de la famille exercent un effet sensible sur le niveau d'instruction des jeunes de 26 ans provenant des milieux urbain et rural. La province de résidence influe de plusieurs façons sur le niveau d'instruction des individus de 26 ans issus d'un milieu rural. En premier lieu, la variable province de résidence influe en soi sur les chances de réussite d'un individu de 26 ans au sein du système d'éducation. En second lieu, la province de résidence interagit également avec le revenu familial pour influer encore davantage sur le niveau d'instruction d'un individu. Pour les individus de 26 ans provenant d'une région métropolitaine de recensement de 100 000 habitants ou plus, l'interaction de la province de résidence et de la valeur du logement influe indirectement sur le niveau d'instruction de l'individu, mais l'effet de la province de résidence en soi est négligeable.

Bien que les résultats concernant les individus de 26 ans paraissent compatibles avec ceux des individus de 18 ans, les implications ne sont pas nécessairement les mêmes. Les associations statistiques entre les variables concernant le niveau d'instruction et le revenu familial sont semblables pour les deux groupes d'âge. Comme on s'y attendait, les individus issus de familles aisées ont tendance à étudier plus longtemps. Cependant, le revenu familial n'entraîne pas les mêmes effets pour les

individus de 18 ans que pour ceux de 26 ans en 1986. La plupart des jeunes de 18 ans sont toujours aux études à temps plein et ne contribuent pas au revenu familial. Par contre, la majorité des individus de 26 ans sont déjà sur le marché du travail et contribuent probablement de façon importante au revenu familial. Par conséquent, la corrélation entre le niveau d'instruction et le revenu familial acquiert une signification différente. La relation dans ce cas est une association statistique (c.-à-d. que le niveau d'instruction et le revenu familial ont tendance à progresser simultanément) sans lien apparent de cause à effet. Étant donné que la majorité des individus de 26 ans ont déjà terminé leur scolarité, les rapports empiriques entre les facteurs socio-économiques et le niveau d'instruction doivent être interprétés avec circonspection.

Les résultats indiquent que la province de résidence, la taille de la famille, le niveau d'instruction et les semaines de travail ont un effet statistiquement significatif sur le revenu du travail. L'effet de la province de résidence, des années de scolarité terminées et des semaines de travail sur le revenu du travail est évidente. La corrélation entre la taille de la famille et le revenu du travail, cependant, est probablement une association statistique plutôt qu'une relation de cause à effet.

Les résultats statistiques montrent que de nombreux facteurs — y compris le revenu familial, la province de résidence, la taille de la famille, la scolarité et l'interaction du revenu familial et de la scolarité — influent de façon systématique sur l'expérience professionnelle de l'individu (c.-à-d. les semaines de travail). Cela s'applique également aux individus provenant des grandes régions métropolitaines de recensement et des régions rurales. Le rapport entre l'expérience professionnelle et l'instruction est tel, toutes choses étant égales par ailleurs, que les individus les plus instruits réussissent toujours mieux sur le marché du travail que les moins instruits.

Effets quantitatifs des facteurs socio-économiques sur le niveau d'instruction

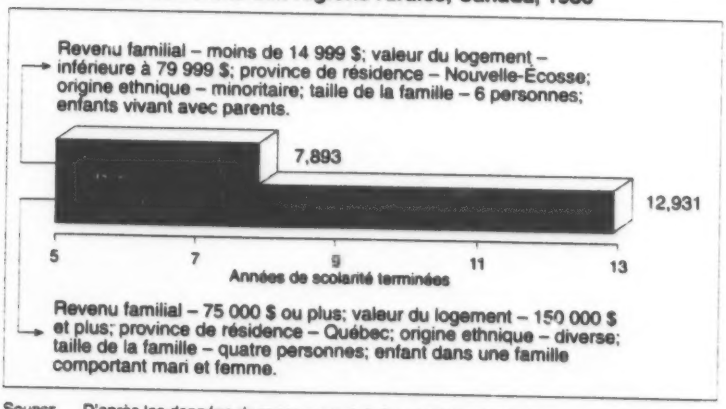
Jusqu'à présent, le rapport n'a fait que présenter séparément les facteurs statistiquement significatifs. Il peut arriver cependant qu'il soit préférable d'avoir des données plus précises. Par exemple, le lecteur peut se demander dans quelle mesure un faible revenu familial et un piètre milieu résidentiel contribuent au faible niveau d'instruction des individus issus de quartiers pauvres, après que les effets de tous les autres facteurs sont simultanément pris en compte. La valeur estimée, qui est l'une des possibilités offertes par le modèle linéaire généralisé de l'analyse de la variance[20], traite directement de ce genre de question. L'examen qui suit porte sur les jeunes de 18 ans provenant de régions rurales, mais les résultats s'appliquent généralement aux autres groupes d'âge et aux régions urbaines.

Pour mesurer l'effet des forces socio-économiques favorables et défavorables sur le niveau d'instruction des individus de 18 ans, nous employons les termes « favorables » et « défavorables » dans un sens hypothétique et explicatif. On dit qu'un jeune de 18 ans provient d'un milieu socio-économique favorable lorsqu'il est issu d'une famille dont le revenu s'établit à 75 000 $ ou plus, laquelle possède sa maison, d'une valeur de 150 000 $ ou plus, lorsqu'il est résident du Québec, provient d'une famille aux origines ethniques diverses (p. ex. un père britannique et une mère française), qui compte quatre membres et compte un mari et une femme. On dit qu'un jeune est issu d'un milieu socio-économique défavorable s'il provient d'une famille à faible revenu (dont le revenu total est inférieur à 14 999 $; vit dans un logement évalué à moins de 79 999 $, habite la Nouvelle-Écosse, provient d'un groupe ethnique minoritaire, est issu d'une famille comptant six membres ou plus, et vit avec des parents. Il va sans dire que dans le monde réel, des combinaisons pires ou meilleures peuvent exister. Ces deux combinaisons socio-économiques extrêmes n'ont été choisies qu'à des fins d'illustration. Elles représentent le meilleur et le pire des mondes possibles, dans les limites de l'échantillon du recensement. Par exemple, il est possible de songer à une meilleure ou à une pire combinaison socio-économique, mais s'il n'existe pas dans les micro-données des individus de 18 ans qui présentent ces caractéristiques socio-économiques, l'effet de ce milieu socio-économique hypothétique est impossible à estimer.

Les individus issus d'un milieu social et économique favorable peuvent s'attendre à avoir terminé 12,9 années de scolarité à l'âge de 18 ans (figure 6). Cette estimation est statistiquement significative au niveau de

Figure 6

Effets des forces socio-économiques sur le niveau d'instruction des individus de 18 ans des régions rurales, Canada, 1986



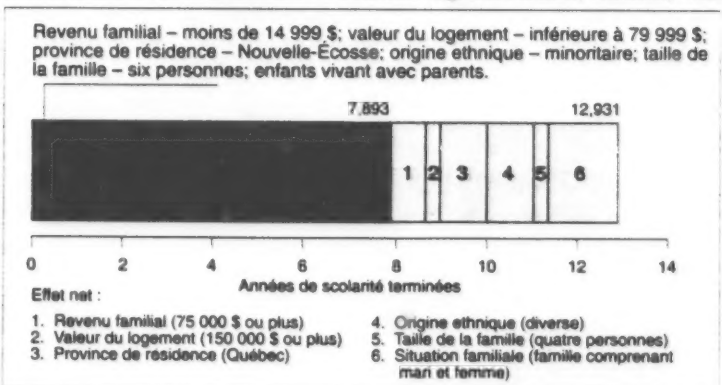
Source D'après les données du recensement de la population de 1986.

0,01 %. Il n'y a pratiquement pas de chance que l'estimation soit un artefact statistique. À l'autre extrémité du spectre, un individu d'un milieu socio-économique défavorable ne peut espérer terminer que 7,9 années de scolarité (figure 6 et tableau A-9). Il faut mentionner que ces estimations se rapportent à un niveau d'instruction moyen. Certains individus pourraient bien sûr réussir mieux ou moins bien selon leurs particularités propres. Par exemple, si un individu est plus doué que les autres membres du groupe, son niveau d'instruction sera probablement supérieur au chiffre estimé.

La figure 7 est un résumé des contributions factorielles des forces socio-économiques. Le graphique à barres subdivisées représente l'effet net de chaque facteur favorable de façon additive et progressive. La première partie du graphique indique que si l'individu est issu d'un milieu socio-économique défavorable, les années de scolarité terminées pour cet individu de 18 ans s'établissent à 7,9 ans de « scolarité terminée ». Cependant, si cet individu provenait d'une famille dont le revenu annuel était de 75 000 $ ou plus, tous les autres facteurs défavorables demeurant inchangés, sa scolarité serait plus élevée. L'augmentation nette du niveau d'instruction est illustrée par le rectangle 1 du graphique à barres subdivisées. De même, les rectangles 2 à 6 représentent l'effet net d'autres facteurs favorables sur la scolarité de l'individu. Dans la vie réelle de même que dans l'analyse statistique, ces facteurs influent simultanément sur l'instruction. Par conséquent, le lecteur doit considérer le graphique uniquement comme un moyen de représentation. Une famille comptant un mari et une femme exerce la plus grande influence sur le niveau d'ins-

Figure 7

Effets des facteurs socio-économiques favorables sur le niveau d'instruction des individus de 18 ans des régions rurales, Canada, 1986



Source D'après les données du recensement de la population de 1986.

truction d'un enfant. Des facteurs tels que le revenu familial, la province de résidence et l'origine ethnique jouent également un rôle considérable dans le niveau d'instruction d'un enfant. En revanche, la taille de la famille et la valeur élevée du logement ont des incidences relativement mineures.

En résumé, l'évaluation a démontré que les facteurs socio-économiques peuvent influer profondément sur la réussite d'une personne au sein du système d'éducation. En fait, selon l'exemple employé ici, l'écart entre des facteurs socio-économiques très favorables et exceptionnellement défavorables peut expliquer environ cinq années additionnelles de scolarité terminée. De toute évidence, si le Canada est intéressé à relever le niveau d'instruction de tous ses citoyens, il lui faudra, en plus de poursuivre sa recherche d'un système d'éducation plus efficace, trouver des façons d'atténuer une partie de l'effet social et économique défavorable sur les jeunes moins favorisés.

Conclusion

La présente analyse a montré que les facteurs sociaux et économiques exercent un effet appréciable sur la scolarité de l'individu. Cette conclusion concorde avec un consensus établi parmi les chercheurs en économie de l'éducation. La présente étude n'apporte aucune conclusion inédite. Au contraire, sa contribution consiste à confirmer empiriquement une croyance populaire. En outre, bien que les observateurs aient reconnu par le passé l'influence des facteurs sociaux et économiques sur l'instruction, ils ont rarement produit des estimations quantitatives des effets d'une série complète de facteurs socio-économiques individuels. La présente étude est peut-être la première à le faire.

L'État a peu de contrôle sur les variables significatives dont on a établi qu'elles influençaient le niveau d'instruction. Par exemple, les décideurs ne sont pas en mesure de modifier la situation économique des familles pauvres, de changer la taille d'une famille ou d'ajouter un conjoint aux familles monoparentales.

A priori, la pauvreté, la taille de la famille et l'absence d'un couple dans une famille ne sont pas directement responsables d'une faible scolarité. Seul le milieu général créé par ces facteurs empêche de parvenir à un niveau élevé de scolarité. Comment l'État peut-il aider les jeunes moins bien nantis à réaliser leur potentiel ? Le programme Head Start des États-Unis a montré qu'il n'existe pas de « solution miracle ». Les programmes préscolaires peuvent aider certains enfants défavorisés pendant un court laps de temps. Cependant, même les partisans les plus enthousiastes du programme Head Start ne s'attendent pas à ce que les

effets bénéfiques du programme se prolongent pendant toute la durée des études d'un enfant, parce que les jeunes de milieux socio-économiques défavorisés ont trop d'occasions d'échouer. Les décideurs ne disposent pas actuellement de solutions véritables à ces problèmes. En Amérique du Nord, les enfants de familles riches sont souvent envoyés dans des établissements scolaires privés pendant leurs années de formation. Ces jeunes privilégiés ont tendance à réussir proportionnellement mieux au sein du système d'éducation que leurs homologues issus de familles pauvres. À priori, leur rendement supérieur doit être attribué au moins en partie au milieu d'étude supérieur des écoles privées. Pourquoi l'État ne peut-il régler le problème socio-économique des enfants moins favorisés en leur offrant une instruction privée gratuite ? La réponse est, bien sûr, que le coût d'un tel programme serait prohibitif. Compte tenu des difficultés financières actuelles, le gouvernement fédéral et les provinces ne peuvent tout simplement pas se permettre une entreprise aussi coûteuse. En outre, même si les pouvoirs publics disposaient des ressources nécessaires à cette fin, il n'est pas évident qu'un grand nombre de parents pauvres choisiraient d'envoyer leurs enfants à des écoles privées gratuites, et personne ne pourrait les y obliger. Un programme d'enseignement gratuit sans élève deviendrait un éléphant blanc. Cela ne réglerait pas la question. Si les décideurs jugent cette solution peu pratique, ils devraient immédiatement consulter les spécialistes en éducation, en psychologie, en criminologie et la police pour trouver une solution de rechange viable.

Appendice

Résumé des tests à une variable

Tableau A-1

Variable dépendante — années de scolarité terminées; échantillon — tous les individus âgés de 18 ans des régions métropolitaines de recensement comptant une population de 100 000 habitants ou plus, 1986

Variable indépendante (facteur de variation)	Valeur F	Niveau de signification[1] (en pourcentage)
Classe de revenu	8,18	0,01
Valeur du logement	17,38	0,01
Province	3,22	0,67
Taille de la famille	6,95	0,01
Situation familiale	4,01	0,13
Classe de revenu — province	1,69	1,70
Classe de revenu — taille de la famille	2,15	1,15

1 Dans l'analyse de type III, le principal effet d'une variable indépendante est corrige des effets des autres variables indépendantes et de leur interaction. Pour de plus amples détails, voir R. J. Freund, R. C. Littell, et P. C. Spector, *SAS System for Linear Models*, Cary, SAS Institute Inc., 1986, et *SAS/STAT User's Guide*, version 6, quatrième ed., Cary, SAS Institute Inc., 1990.

SOURCE D'après des données du recensement de la population de 1986.

Tableau A-2

Variable dépendante — Années de scolarité terminées; échantillon — tous les individus âgés de 18 ans habitant dans des villages, des villes et des régions rurales, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Classe de revenu	16,63	0,01
Valeur du logement	6,75	0,01
Province	14,78	0,01
Origine ethnique	16,67	0,01
Taille de la famille	8,28	0,01
Situation familiale	12,68	0,01

Tableau A-3

Variable dépendante — années de scolarité terminées; échantillon — tous les individus âgés de 26 ans habitant dans des régions métropolitaines de recensement ayant une population de 100 000 habitants ou plus, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Classe de revenu	55,80	0,01
Valeur du logement	5,21	0,03
Origine ethnique	6,40	0,01
Taille de la famille	26,53	0,01
Province — valeur du logement	1,71	1,68

Tableau A-4

Variable dépendante — revenu salarial en 1985; échantillon — tous les individus âgés de 26 ans habitant dans des régions métropolitaines de recensement ayant une population de 100 000 habitants ou plus, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Province	6,41	0,01
Taille de la famille	4,13	0,02
Classe d'instruction	14,17	0,01
Classe des semaines de travail	556,77	0,01

Tableau A-5

Variable dépendante — semaines de travail en 1985; échantillon — tous les individus âgés de 26 ans habitant dans des régions métropolitaines de recensement ayant une population de 100 000 habitants ou plus, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Classe de revenu	133,57	0,01
Province	8,65	0,01
Taille de la famille	71,34	0,01
Classe d'instruction	11,06	0,01
Classe de revenu — classe d'instruction	2,59	0,07

Tableau A-6

Variable dépendante — années de scolarité terminées; échantillon — tous les individus âgés de 26 ans habitant dans des villages, des villes et des régions rurales, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Classe de revenu	25,98	0,01
Province	5,51	0,01
Origine ethnique	10,15	0,01
Taille de la famille	44,42	0,01
Classe de revenu — province	1,56	0,75

Tableau A-7

Variable dépendante — revenu salarial en 1985; échantillon — tous les individus âgés de 26 ans habitant dans des villages, des villes et des régions rurales, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Province	6,25	0,01
Taille de la famille	5,59	0,02
Classe d'instruction	16,06	0,01
Classe de semaines de travail	514,24	0,01

Tableau A-8

Variable dépendante — semaines de travail en 1985; échantillon — tous les individus âgés de 26 ans habitant dans des villages, des villes et des régions rurales, Canada, 1986

Variable indépendante (Facteurs de variation)	Valeur F	Niveau de signification (en pourcentage)
Classe de revenu	85,09	0,01
Province	7,55	0,01
Taille de la famille	73,20	0,01
Classe d'instruction	15,66	0,01
Classe de revenu — classe d'instruction	1,98	1,34

Tableau A-9

Effets quantitatifs des facteurs socio-économiques sur l'instruction, Canada, 1986

Individus de 18 ans des centres urbains

Influence défavorable

Revenu familial — 14 999 $ ou moins en 1985
Valeur du logement — 79 999 $ ou moins
Taille de la famille — six personnes ou plus par famille
Structure familiale — enfants vivant avec des parents

Estimation des années de scolarité terminées — 9,610
(significatif statistiquement à 0,01 pour cent)

Influence favorable

Revenu familial — 75 000 $ ou plus en 1985
Valeur du logement — 150 000 $ ou plus
Taille de la famille — 4 personnes par famille
Structure familiale — enfants vivant dans une famille comportant mari et femme

Estimation des années de scolarité terminées — 12,422
(statistiquement significatif à 0,01 pour cent)

Individus de 18 ans des régions rurales

Influence défavorable

Revenu familial — 14 199 $ ou moins en 1985
Valeur du logement — 79 999 $ ou moins
Province de résidence — Nouvelle-Écosse
Origine ethnique — minoritaire
Taille de la famille — six personnes ou plus par famille
Structure familiale — enfants vivant avec des parents

Estimation des années de scolarité terminées — 7,893
(statistiquement significatif à 0,01 pour cent)

Influence favorable

Revenu familial — 75 000 $ ou plus en 1985
Valeur du logement — 150 000 $ ou plus
Province de résidence — Québec
Origine ethnique — diverse (p. ex. Britannique ou autres)
Taille de la famille — 4 personnes par famille
Structure familiale — enfants vivant dans une famille comportant mari et femme

Estimation des années de scolarité terminées — 12,931
(statistiquement significatif à 0,01 pour cent)

Notes

1 Voir Gamme Coleman *et al.*, *Equality of Educational Opportunity*, Washington (D.C.), département de la Santé, de l'Éducation et du Bien-être social, 1966.

2 Voir James Coleman *et al.*, *Equality of Educational Opportunity*; et Burt S. Barnow et Glen G. Cain, « A reanalysis of the effect of Head Start on cognitive development: Methodology and empirical findings », *Journal of Human Resources*, printemps 1977, vol. 12, n° 2, p. 177-197.

3 W. S. Barnett, « Benefit-cost analysis of the Perry Preschool Program and its long-term effects », *Educational Evaluation and Policy Analysis 1985*, vol. 4, n° 7, p. 333-342.

4 Voir Maisy Cheng *et al.*, *The Every Secondary Student Survey*, automne 1987, Toronto Board of Education, juin 1989.

5 Voir Coleman *et al.*, *Equality of Educational Opportunity*.

6 Voir, par exemple, Barnow et Cain, « A reanalysis of the effect of Head Start ».

7 Barnett, « Benefit-cost analysis of the Perry Preschool Program ».

8 Richard J. Murnane, « What do effective teachers of inner-city children have in common ? », *Social Science Research*, mars 1981, vol. 10, n° 1.

9 Anita A. Summers et Barbara L. Wolfe, « Do schools make a difference ? », *The America Economic Review*, septembre 1977, vol. 67, n° 4.

10 Richard J. Murnane *et al.*, « Home resources and children's achievement », *Review of Economics and Statistics*, août 1981, vol. 63, n° 3, p. 369-377.

11 Voir, par exemple, Summers et Wolfe, « Do schools make a difference ? »

12 Maisy Cheng *et al.*, *The Every Secondary Student Survey*.

13 Voir, par exemple, Murnane *et al.*, « Home resources and children's achievement ».

14 Le fichier de micro-données du recensement comprend des données distinctes pour les 10 provinces, le Yukon et les Territoires du Nord-Ouest étant combinés. Par conséquent, 10 variables auxiliaires suffisent pour saisir les effets des provinces. L'effet du territoire omis est inclus dans le terme constant de l'équation d'estimation.

15 L'analyse de la variance est une théorie bien établie dans l'histoire de la statistique. Le lecteur trouvera facilement de bons ouvrages sur la question. Voir, par exemple, S. R. Searle, *Linear Models*, New York, John Wiley & Sons, 1971, et F. A. Graybill, *Theory and Application of the Linear Model*, Belmont (Cal.), Wadsworth Publishing, 1976.

16 Ce phénomène statistique est mal connu de ceux qui utilisent la méthode de la régression multiple. Pour de plus amples renseignements, voir F. A. Graybill, *Theory and Application of the Linear Model*; R. J. Freund *et al.*, *SAS System for Linear Models*, Cary, SAS Institute, 1986; et *SAS/2STAT User's Guide*, Version 6, quatrième éd., Cary, SAS Institute, 1990.

17 À ce propos, nous pourrions souligner que les données du recensement utilisées dans la présente étude sont toujours sous forme non équilibrée. Le modèle d'analyse linéaire généralisé de la variance nous permet de résoudre correctement et efficacement ce problème de données. Voir *SAS/STAT User's Guide*.

18 Nous n'avons pas complètement négligé les individus âgés de 15 à 17 ans et de 19 à 25 ans. Pour réduire la somme des travaux statistiques, nous avons partagé les jeunes en trois groupes d'âge — 15–17, 19–22, et 23–25. Cette méthode nous a permis d'étudier le comportement de ces individus de façon économique, sans soulever trop de problèmes d'agrégation relatifs à l'âge. Les résultats obtenus pour les individus de 15 à 17 ans et de 19 à 22 ans correspondent d'assez près à ceux des individus de 18 ans, et les résultats des individus âgés de 23 à 25 ans sont semblables à ceux des individus de 26 ans. Ces résultats peuvent être obtenus de l'auteur, sur demande.

19 Pour de plus amples détails sur cette question, voir Freund *et al.*, *SAS System for Linear Models;* et *SAS/STAT User's Guide*.

20 Voir *SAS/STAT User's Guide*.